

Online Research @ Cardiff

This is an Open Access document downloaded from ORCA, Cardiff University's institutional repository: <https://orca.cardiff.ac.uk/id/eprint/119177/>

This is the author's version of a work that was submitted to / accepted for publication.

Citation for final published version:

Han, Jiaying, Permentier, Hjalmar, Bischoff, Rainer, Groothuis, Geny, Casini, Angela ORCID: <https://orcid.org/0000-0003-1599-9542> and Horvatovich, Péter 2019. Imaging of protein distribution in tissues using mass spectrometry: An interdisciplinary challenge. Trends in Analytical Chemistry 112 , pp. 13-28. 10.1016/j.trac.2018.12.016 file

Publishers page: <http://dx.doi.org/10.1016/j.trac.2018.12.016>
<<http://dx.doi.org/10.1016/j.trac.2018.12.016>>

Please note:

Changes made as a result of publishing processes such as copy-editing, formatting and page numbers may not be reflected in this version. For the definitive version of this publication, please refer to the published source. You are advised to consult the publisher's version if you wish to cite this paper.

This version is being made available in accordance with publisher policies.

See

<http://orca.cf.ac.uk/policies.html> for usage policies. Copyright and moral rights for publications made available in ORCA are retained by the copyright holders.



Imaging of protein distribution in tissues using mass spectrometry: an interdisciplinary challenge

Jiaying Han^{1,2}, Hjalmar Permentier¹, Rainer Bischoff^d, Geny Groothuis², Angela Casini^{2,3}, Péter Horvatovich^{1}*

¹Analytical Biochemistry, Groningen Research Institute of Pharmacy, University of Groningen, Antonius Deusinglaan 1, 9713 AV Groningen, The Netherlands

²Pharmacokinetics, Toxicology and Targeting, Groningen Research Institute of Pharmacy, University of Groningen, Antonius Deusinglaan 1, 9713 AV Groningen, The Netherlands

³School of Chemistry, Cardiff University, Main Building, Park Place, CF10 3AT Cardiff, United Kingdom

Abbreviations

AP	atmospheric pressure
CID	collision-induced dissociation
CT	Computed Tomography
DCE	Datacube Explorer
DESI	desorption electrospray ionization
ECD	electron capture dissociation
ETD	electron transfer dissociation
ESI	electrospray
FAIMS	high field asymmetric waveform ion mobility
FFPE	formaldehyde-fixed and paraffin-embedded
FTICR	Fourier transform ion cyclotron resonance
H&E	hematoxylin and eosin
HCD	higher-energy collision induced dissociation
IHC	immunohistochemistry
IMS	ion mobility separation

28	IRMPD	infrared multiphoton dissociation
29	ISD	in-source decay
30	LAESI	laser ablation electrospray ionization
31	LA-ICP MSI	laser ablation inductively coupled plasma mass spectrometry imaging
32	LC/ESI-MS/MS	liquid chromatography electrospray ionization tandem mass spectrometry
33	LC-MS	liquid chromatography mass spectrometry
34	LID	laser-induced dissociation
35	MALDI	matrix assisted laser desorption ionization
36	ME-SIMS	matrix enhanced secondary ion mass spectrometry
37	MITICS	MALDI imaging team imaging computing system
38	MS/MS	tandem mass spectrometry
39	MSI	mass spectrometry imaging
40	NHS	N-hydroxysuccinimide
41	PC	photocleavable
42	3-SBA	3-sulfobenzoic acid
43	scFv	single chain variable fragment
44	SIMS	secondary ion mass spectrometry
45	SMALDI	scanning microprobe matrix assisted laser desorption ionization
46	4-SPITC	4-sulphophenyl isothiocyanate
47	TAMSIM	targeted multiplex mass spectrometry imaging
48	TMPP	N-succinimidylloxycarbonylmethyl)-tris(2,4,6-trimethoxyphenyl)phosphonium
49		bromide
50	TOF	time-of-flight
51	UVPD	ultraviolet photodissociation
52		

53 **Abstract**

54 The recent development of mass spectrometry imaging (MSI) technology allowed to obtain highly
55 detailed images of the spatial distribution of proteins in tissue at high spatial resolution reaching cell
56 dimensions, high target specificity and a large dynamic concentration range. This review focusses on
57 the development of two main MSI principles, targeted and untargeted detection of protein distribution
58 in tissue samples, with special emphasis on the improvements in analyzed mass range and spatial
59 resolution over the last 10 years. Untargeted MSI of *in situ* digested proteins with matrix-assisted
60 laser desorption ionization is the most widely used approach, but targeted protein MSI technologies
61 using laser ablation inductively coupled plasma (LA-ICP) and photocleavable mass tag chemical
62 labeling strategies are gaining momentum. Moreover, this review also provides an overview of the
63 effect of sample preparation on image quality and the bioinformatic challenge to identify proteins and
64 quantify their distribution in complex MSI data.

65

66 **Introduction**

67 Proteins participate actively in biological events and fulfill a wide range of molecular functions, such
68 as substrate transport, cellular signaling, catalysis of metabolic reactions, and regulation of DNA
69 replication and transcription events. Protein expression changes may indicate the presence and
70 severity of a disease, and can be used to identify disease onset at an early stage, providing better
71 treatment options for patients. Tissues are particularly important samples in clinical research, because
72 they contain rich information on morphologic, metabolomic and proteomic changes related to
73 biological events and disease pathology^{1,2}. The imaging of protein distribution in tissues can provide
74 new insights into the molecular mechanisms of diseases and the normal function of cells and tissues,
75 as well as of aging processes. The spatial distribution of proteins in tissue samples provides
76 information that is complementary to the relative and absolute concentration information obtained
77 with commonly applied high-throughput molecular profiling omics approaches, such as liquid
78 chromatography mass spectrometry (LC-MS)-based proteomics and metabolomics.

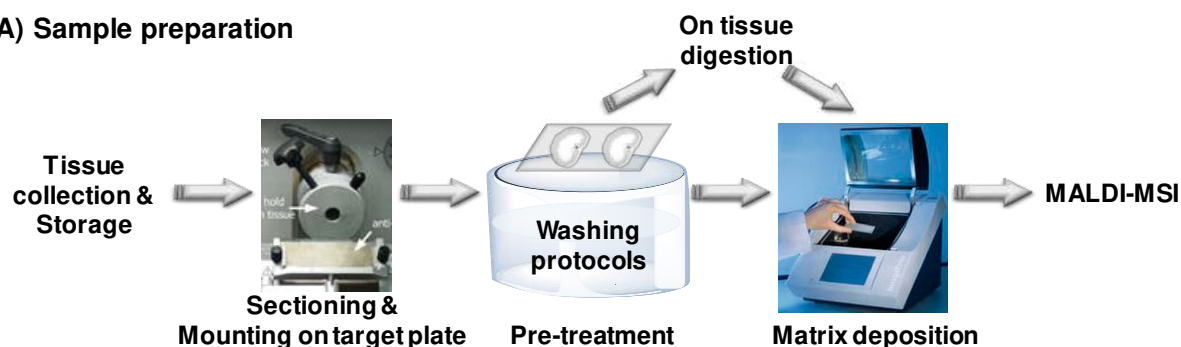
79 In order to obtain an image from a complex tissue specimen, several non-invasive imaging approaches
80 have been developed such as radiography (X-ray, Computed Tomography (CT))³, ultrasonography
81 (USG)⁴, positron emission tomography (PET)⁵ and magnetic resonance imaging (MRI)⁶ making use
82 of different measurable physicochemical properties such as emitted/reflected light, particles (e.g.
83 positrons) and ultrasound. These approaches have contributed significantly to the visualization of
84 biological processes and many of them are applied routinely in clinical diagnostics. While many
85 commonly used “non-invasive” (not requiring tissue sampling from patient) imaging technologies,
86 such as CT and X-ray radiography, and “invasive” (requiring tissue sampling from patients) imaging
87 technologies, including those based on ultraviolet-visible (UV-VIS) and fluorescence spectroscopy,
88 are applied to provide high-quality images from tissues, this information cannot always be
89 straightforwardly translated into an image reflecting the spatial distribution of individual analytes (e.g.
90 proteins). Immunostaining in combination with optical or fluorescence imaging can provide signals
91 from specific proteins by visualizing the distribution of antibody-antigen pairs in tissue. However,
92 images acquired with UV-VIS, fluorescence and radiography^{7,8} usually provide spatial distribution
93 for only a limited number of proteins in a single experiment. In addition, most methods require *a*
94 *priori* knowledge of the target molecules, which prevents their use as a hypothesis-free discovery and
95 hypothesis-generating tool. Some imaging technologies measure the physicochemical properties of
96 an ensemble of compounds, with spatial localisation in tissue such as nuclear magnetic resonance
97 spectroscopy, or common UV-VIS microscopy^{9,10}, therefore, only inferring the presence of some
98 compounds or classes of compounds. In this context, mass spectrometry imaging (MSI) is a powerful
99 alternative, which circumvents some of these limitations.

100 In fact, MSI takes full advantage of the high chemical specificity of mass spectrometry and allow to
101 quantify the spatial distribution of hundreds of individual molecules in tissues in a single
102 measurement, without the need for labels or prior knowledge of the analytes. In addition, MSI
103 technology allows to detect in one experiment multiple compounds which do not ionize well or are
104 in low abundance, using reagents specifically targeting these compounds. Nowadays, there are several
105 MSI approaches, which differ in the way that compounds are desorbed into the gas phase and ionized
106 for sampling into the mass spectrometer, including secondary ion mass spectrometry (SIMS), MALDI
107 MSI, LA-ICP MSI, desorption electrospray ionization (DESI), rapid evaporative ionization mass
108 spectrometry (REIMS)¹¹, direct analysis in real time (DART)¹² and easy ambient sonic-spray
109 ionization (EASI)¹³. Thus, the unique features of MSI to sample compounds directly allows analysis
110 of many types of (bio)molecules such as proteins, metabolites and drugs, and provides potential for
111 a wide range of research applications. Examples of these applications include approaches which
112 provide new insight into normal and disease-related molecular processes^{14,15,16}, enable disease
113 prognosis and prediction of response to therapy, allow to obtain the distribution of a drug in its intact
114 form and its metabolites in tissue^{17–20}, or provide classification of tissues based on molecular
115 information and reveal details of microbiome molecular communication²¹.

116 This review focuses on state-of-the-art MSI approaches used to determine protein distribution in
117 tissues. In details, the manuscript discusses the technical aspects of protein MSI, such as sample
118 preparation, protein desorption in the gas-phase and ionization, spatial resolution and measured
119 dynamic concentration range, and presents in detail various MSI approaches for targeted and
120 untargeted detection of protein distributions in tissue samples. This includes the most commonly used
121 untargeted protein imaging MSI using MALDI, and other ion generation and sampling approaches
122 such as LA-ICP MSI, and targeted protein MSI using chemical labeling with photocleavable mass
123 tags (e.g. Tag-Mass)^{22–25}. One section discusses the data processing and interpretation challenge
124 related to protein MSI. The review ends with a discussion of the possible future directions of MSI
125 methodologies for protein distribution analysis in tissue samples.

1 Main steps of protein distribution analysis in tissue using mass spectrometry imaging

(A) Sample preparation



(B) Mass spectrometry analysis of local peptides/protein content

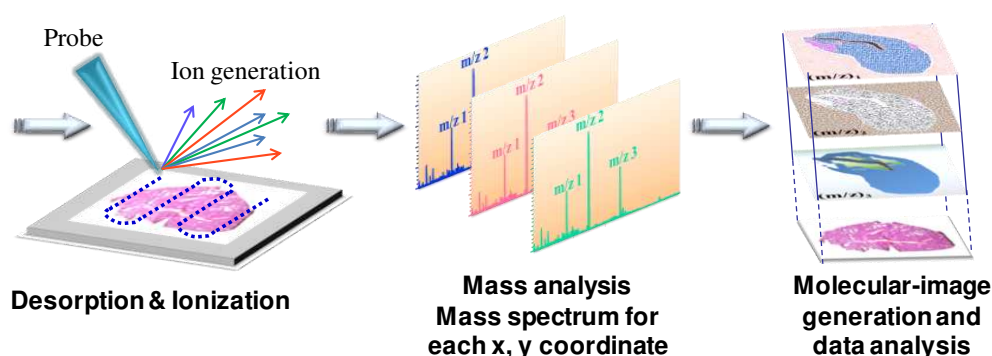


Figure 1. An example of mass spectrometry imaging (MSI) using a MALDI interface, which is a commonly used workflow for peptide/protein distribution analysis in tissue including tissue sectioning and sample preparation (A) and data acquisition and evaluation (B).

MSI of protein distribution in tissue samples consists of three main steps: sample preparation, data acquisition and data (pre-)processing and interpretation (**Figure 1**). The sample preparation protocols have a crucial impact on the quality of the MSI process. Sample preparation has the goal to facilitate the desorption into the gas-phase and the ionisation of proteins or peptides obtained after trypsin digestion, while keeping protein diffusion to a minimum and maintaining the original spatial distribution of proteins. These two aims are conflicting, and their balance plays a crucial role in the quality of the obtained MSI image. The mass spectrometer interface determines the desorption in the gas-phase, ionization and ion sampling efficiency, the speed of sampling and the area from which the ions are sampled. The latter property determines the theoretical spatial resolution of the MSI image. Theoretical spatial resolution can only be reached if sample preparation ensures that local protein abundance is maintained in the tissue to be imaged. The mass analyzer and acquisition parameters determine the speed of data acquisition, the type of registered spectra (with or without fragmentation), the measured dynamic range and the resolution of the acquired mass spectra. Bioinformatics solutions to pre-process and analyze MSI data form an important part of protein MSI workflows and have the

goal to interpret the large amount of collected protein distribution information together with other metadata such as a histology image with anatomical annotation by an expert pathologist^{26,27}.

1.1 Tissue sample preparation

Tissue sample preparation is probably the most critical step to obtain optimal sensitivity, reproducibility and spatial resolution of the protein distribution in an MSI experiments^{28,29}. Inappropriate sample preparation leading to protein degradation, signal interference by non-target chemical species, alteration of the original protein distribution, or low ion sampling efficiency due to insulating properties of tissue have a negative effect on the quality of the acquired MSI data. Normally, tissue sample preparation involves organ harvesting and tissue sectioning (**Figure 1**). In order to avoid delocalization and degradation of proteins, it is essential to handle tissues correctly starting with the surgical removal process. After removal of the tissue from the body, tissue samples should be immediately snap-frozen in 2-methyl-butane (isopentane) and stored at -80 °C until use. For MSI of proteins, fresh frozen tissue is preferred over alcohol-preserved, or formaldehyde-fixated and paraffin-embedded (FFPE) tissue sections, because of the covalent crosslinking of proteins in FFPE sections or precipitated proteins in alcohol-preserved sections, although recently the antigen retrieval strategy has been suggested to overcome the protein crosslinks in FFPE sections^{30,31}. In all cases, tissue sections with a thickness of approximately 10 µm are prepared with a microcryotome. It is important to place the frozen tissue sections on sample plates or conductive glass slides without scratches, rips or tears. Once the section (at this point still frozen) is in position, it is thaw-mounted by warming the bottom of the sample plate for macroscopic drying of the section which usually takes 20-30 seconds. Freeze-drying of tissue sections is an often performed operation, however many researchers omit this step from their tissue preparation pipelines without issue³². The sample plate and tissue section are quickly warmed together, resulting in no loss of water-soluble proteins nor translocation of the proteins due to diffusion in the liquid state³³.

Biological tissues contain numerous chemical species over a wide range of concentrations, and more abundant and/or easier ionizable species such as lipids can suppress the detection of less abundant species due to charge competition of compounds during ionization. For instance, salts and lipids³⁴ will interfere with MALDI MSI of proteins or peptides. To partially overcome these problems, tissue-washing procedures have been introduced prior to matrix deposition when using the MALDI MSI method. These washing protocols vary greatly depending on the target analytes. Ideally, all of the unwanted chemical species should be removed from the tissue while maintaining tissue morphology and not disturbing the original spatial distribution of soluble proteins. Assessment of all tissue-washing steps is necessary since each one may lead to some degree of disruption of the original spatial distribution of analytes in the tissue.

181 Matrix application is required for some of the approaches such as MALDI MSI or matrix enhanced
182 secondary ion mass spectrometry (ME-SIMS) MSI. The most widely used technique for MSI is
183 MALDI, for which the reproducibility of the ionization process is still a challenge and the MS
184 acquisition parameters are difficult to optimize. Matrices such as 3,5-dimethoxy-4-hydroxycinnamic
185 acid (sinapinic acid) and α -cyano-4-hydroxycinnamic acid are generally used to promote ionization
186 and prevent degradation of target compounds by the probe beam (laser) energy. Generally, ion signal
187 intensities in MALDI-MS are strongly influenced by the choice of matrix compound and the matrix
188 preparation and deposition procedure, which determines the size distribution of the matrix crystals.
189 Matrix crystal size is the most important parameter, which influences the ionization efficiency and
190 reproducibility of desorption in the gas-phase of compounds. The goal of the procedure is to obtain a
191 homogeneous distribution and uniformly small crystal sizes of matrix for optimal performance³⁵.
192 Several matrix application and drying cycles can be performed until an optimal matrix thickness with
193 high quality and homogeneity is achieved. There are several methods by which a homogeneous matrix
194 layer can be applied, such as spraying, solvent free dusting or coating by sublimation³⁶, and manual
195 or robotic spotting^{37,38}. Manual spraying is an often used, simple approach for matrix application
196 which works well in the hands of an experienced operator, without requiring sophisticated
197 instrumentation. However, automatic deposition provides a more homogenous matrix layer and
198 improved reproducibility enhancing the imaging performance. The review by Goodwin on commonly
199 used matrix and matrix applications approaches for MSI provides more details on this topic²⁹.

200 In MALDI MSI, proteins are measured with two approaches: either in their intact form, where smaller
201 proteins are easier to measure than large proteins, or after digestion using a protease such as trypsin,
202 which has the advantage that there is no limit with respect to protein size. Mass spectrometers with
203 higher mass resolution allow to achieve better mass accuracy which improves identification of
204 peptides and proteins. Moreover, since peptides are easier to detect and identify, this facilitates
205 subsequent identification of proteins and their post-translational modifications. These two approaches
206 are discussed in detail in sections 2.1.3 and 2.1.4.

207 **1.2 Desorption and ionization of peptides and proteins from tissue**

208 The choice of desorption (extraction into the gas-phase) and ionization technique has an important
209 influence on the spatial resolution of the obtained MSI image and on the detected compound profile
210 (**Figure 2**). SIMS is using high-energy primary ion beams of ionized noble gas, oxygen, fullerene or
211 SF₆, to generate and to sputter secondary ions from sample surface. SIMS was introduced to MSI in
212 the 1960s and was developed to detect atoms or small fragments of vitamins, pharmaceuticals, lipids
213 and peptides in tissue and cells³⁹⁻⁴¹. SIMS was applied to obtain information on elemental, isotopic
214 and molecular composition of the upper atomic layers of the scanned sample^{42,43}. It has the primary

215 advantage of achieving a high spatial resolution (< 100 nm or even ≤ 20 nm), which cannot be
216 achieved with MALDI, LDI or DESI interfaces^{44,45,54–57,46–53}. However, SIMS suffers from severe in-
217 source fragmentation of biomolecules due to excessively hard ionization, which results in impaired
218 identification of target analytes. The lower sensitivity of SIMS-MSI in comparison to MALDI MSI
219 in detecting peptides and proteins was reported in several studies⁵⁸.

220 DESI is an ambient ionization technique developed by Zoltán Takáts, Graham Cooks and their
221 coworkers in 2004 at Purdue University^{59,60}. In this method, a fast, nebulized electrospray gas jet
222 transports charged microdroplets of an eluent to impact the surface of the sample and to carry away
223 ionized molecules. The approach requires no or limited sample preparation effort and allows simple
224 MSI under ambient conditions preventing change in tissue slice shape. Furthermore, DESI is a spray-
225 based soft ionization technique with an average internal energy deposition of ~ 2 eV, which is similar
226 to the internal energy deposition of electrospray (ESI)⁶¹. Thus, DESI yields minimal fragmentation
227 of large molecules compared to the excessive fragmentation of SIMS⁶² and avoids interference with
228 the matrix compounds, such as observed in MALDI. DESI MSI and other variants, such as nano-
229 DESI, have been used for imaging compounds in the low mass region below 1000 Da with a high
230 spatial resolution (approximately 10 μm), as shown for metabolites in leaf tissues or drugs (e.g.
231 clozapine) distribution in animal tissue sections and microbiome sampling^{18,63–67}. The spot size and
232 spatial resolution in (nano)-DESI-MSI – amongst other parameters – depend on the capillary diameter,
233 angle of spray incidence and the tip-to-surface distance, which can be difficult to optimize⁶⁸. (nano)-
234 DESI MSI suffers from the limitation of a much lower spatial resolution compared to SIMS and
235 MALDI, which is for (nano)-DESI typically around 100 μm or upwards in imaging of peptides or
236 proteins^{69–72}. Recently, Garza *et al.* presented a DESI-high field asymmetric waveform ion mobility
237 (FAIMS) device for protein mass spectrometry imaging and reported to simultaneously detect lipids
238 and intact protein forms in mouse kidney, mouse brain, and human ovarian and breast tissue samples⁷³.

239 Another ambient ionization method is laser ablation electrospray ionization (LAESI)^{74,75}, which was
240 introduced by Vertes *et al.* in 2007 and combines laser ablation with a mid-infrared laser and
241 electrospray ionization, where the latter serves to ionize the laser ablated compounds⁷⁴. LAESI does
242 not require complex sample preparation for MSI of peptides or proteins. However, it also suffers from
243 low lateral resolution, which does not allow detailed (sub)cellular imaging.

244 It is necessary to find a technology to overcome all of the above-mentioned issues that can be used
245 for imaging protein distributions in tissue samples. In this context, currently three MSI approaches
246 are used: (1) untargeted MSI of proteins using MALDI, (2) targeted MSI of proteins based on
247 detecting metals ions in their active sites or structural domains or metal ions coupled to antibodies
248 using LA-ICP MSI such as used in mass cytometry, and (3) targeted MSI of proteins using chemical

249 labeling, where the chemical label consists of a protein targeting affinity moiety (antibody, affirmers,
250 activity probes) coupled with photocleavable (PC) mass tags, where mass tag labels are released and
251 measured with MALDI or LDI.

252 MALDI was the first MS-based method for imaging intact proteins in a human glioma⁷⁶ and is
253 currently by far the most commonly used untargeted MSI approach for imaging protein
254 distribution^{77,78}. The first application of MALDI MSI in mapping peptides and proteins in biological
255 samples was developed by the groups of Bernard Spengler (1994)⁷⁹ and Richard Caprioli (1997)⁸⁰.
256 MALDI MSI has since become a mature technology to determine the distribution of proteins over a
257 large mass range from hundreds of Da to beyond 100 kDa with little or no fragmentation of the
258 original protein^{81,82,83–85}. During the last decade, MALDI imaging has been further improved, with
259 respect to detection sensitivity and spatial resolution^{86,87–91}. Current methods can reach a spatial
260 resolution of 10-20 μm ^{84,92}, a value that is limited by the crystal size distribution of the matrix, and
261 therefore does not reach the typical spatial resolution of 100-250 nm of (nano)SIMS. In a typical
262 MALDI MSI interface, ions are formed under vacuum, which constraints the choice of matrix, and
263 may change tissue section morphology. To overcome these problems, atmospheric pressure MALDI
264 (AP-MALDI) ion sources have been developed for MSI applications, where ions are generated at
265 ambient conditions and transferred into the vacuum of the mass analyzer using methods similar to
266 those developed for introduction of ions generated via ESI. AP-MALDI MSI using IR or UV lasers
267 provides high spatial resolution (1.4 μm) in mapping small biomolecules, such as metabolites, lipids,
268 peptides and carbohydrates, but has so far not been applied for protein MSI^{93–97}. In addition, lower
269 sensitivities are observed with AP-MALDI than with vacuum MALDI sources in analyzing plant
270 metabolites⁹⁸.

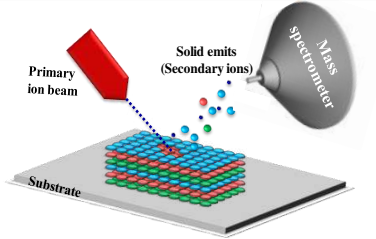
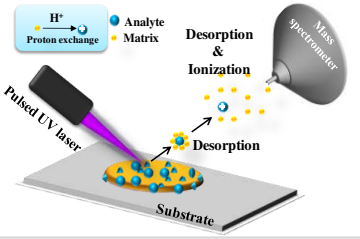
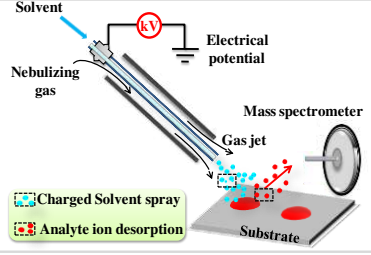
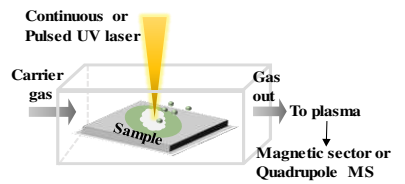
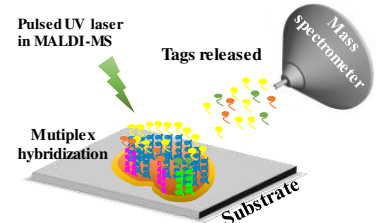
Technique	Schematic	Mass range (Da)	Resolution (μm)	Benefits	Challenges
SIMS		Dynamic SIMS 1-300 Static SIMS 100-1,500	Dynamic SIMS 0.03-0.5 Static SIMS 0.5-50	- high lateral spatial resolution - deep tissue profiling - low LOD* - low depleted material - quantitative imaging (dynamic SIMS) [§]	- source-induced fragmentation - small sampling area - need for conductive mounting surface (need for insulation) - stability of sample under UHV
MALDI		300-200,000	20-500	- wide mass range - AP* and IP* ion sources available - tolerance for large sampling area	- matrix interference - need for conductive mounting surface (needed for insulation) - inaccurate quantitative imaging - inability to measure low m/z (< 300)
DESI		100-200,000	100-1,000 NanoDESI 10-1,000	- soft ionization - real-time imaging of dynamic processes - allow ionization under AP	- low lateral spatial resolution - inaccurate quantitative imaging - need for conductive mounting surface (need for insulation) - high gas pressure altering sample surface - information on elemental and atomic-cluster
LA-ICP		trace elements and specific isotope distribution of metal ions in proteins	10-100 μm	- low limit of detection ($< \mu\text{g g}^{-1}$) - wide dynamic quantification range	- elemental fractionation - interfering matrix effects
Mass tag		50-2000 - detection of large membrane proteins	10-100 μm	- no need for matrix - no mass range limitation	- requirement of expensive reagent - inaccurate quantification due to losses during chemical reaction and instability of the PC-linker during sample treatment

Figure 2. The main characteristics of desorption (extraction in the gas-phase) and ionization interfaces used for mass spectrometry imaging. Abbreviations: LOD: Limit of detection; AP: Atmosphere pressure; IP: Intermediate pressure. UHV: Ultra-high vacuum. *Static SIMS MSI detection of intact molecules above 1,500 Da from biological samples is rarely reported owing to source-induced fragmentation and high LOD for peptides and proteins. LAESI combining DESI and LA for desorption-ionisation was not included in the figure.

275 Other laser irradiation-based desorption/ionization MSI interfaces have been used in protein MSI
276 besides those mentioned above, such as matrix-assisted laser desorption electrospray ionization
277 (MALDESI)⁹⁹, and infrared laser desorption electrospray ionization (IR-LDESI)¹⁰⁰. LA-ICP MSI is
278 another popular approach used for imaging trace elements (e.g. metals and metalloids) in biological
279 materials with a spatial resolution ranging from 200 μm down to 10 μm for a wide range of
280 applications, among them visualization of metal-containing proteins^{101–103}.

281 **1.3 Data processing and analysis**

282 **1.3.1 Spatial resolution in MSI**

283 Spatial resolution is a key parameter to assess the performance of MSI. Spatial resolution is defined
284 in the imaging field as the ability to distinguish two data points with different information content
285 separated in units of distance such as mm or μm . Current MSI technology is able to provide data at
286 low and submicron resolution, however, the spatial resolutions of 10-50 nm¹⁰⁴ achieved by super-
287 resolution imaging is still not achievable. The term spatial resolution is used in multiple contexts,
288 which often leads to confusion. The concept and definition of spatial resolution in the context of MSI
289 is provided here. In general, a tissue is a three-dimensional (3D) compartment, whose MS imaging
290 also provides 3D data, with three coordinate dimensions in tissue and one mass spectrum for each
291 spatial coordinate. A general imaging approach such as MRI, PET or CT collects information on the
292 entire 3D volume of data and in this context two types of resolution are defined: in the axial and the
293 lateral dimension. Axial (longitudinal, azimuthal, range, radial, and depth¹⁰⁵) resolution is defined in
294 parallel to the probe beam of electrons, ions, or photons and defines the ability to distinguish
295 structures at various depths of the sample with respect to the tissue surface¹⁰⁶.

296 Conversely, lateral resolution is defined perpendicular to the probe beam and defines the ability to
297 distinguish structures which lie close to each other side by side, as individual objects. Lateral
298 resolution is affected by the width of the beam, the difference between two adjacent coordinates (step
299 size of sampling) at the tissue surface, but also depends on the depth of imaging i.e. the distance that
300 the beam penetrates the tissue surface, since compounds are sampled from a tissue volume reached
301 effectively by the sampling beam. Wider beams typically scatter in the tissue section and therefore,
302 lateral resolution is improved by using narrower beams and beams that do not penetrate the tissue to
303 great depths¹⁰⁷. MSI is a surface scanning technology, with a low penetration depth into the sample,
304 which is generally applied to a tissue section of a few μm thickness. Therefore, lateral spatial
305 resolution in the plane of the tissue section is the important resolution parameter and this is the
306 definition of spatial resolution used in this review. MSI techniques which acquire data from 3D

307 samples achieve this by merging mass spectrometry ion intensity data from adjacent tissue
308 sections^{108–112}.

309 Spot size, pixel size and step size are important terms to describe the lateral resolution obtained in
310 MSI. Spot size refers to the focus area of the probe beam (laser pulse, ion beam, etc.)¹¹³, which has
311 two definitions; one is based on a Gaussian distribution model of the beam intensity, or irradiance,
312 across its standard deviation, while the second definition expresses the width of the beam at half-
313 intensity^{114,115}. Pixel size refers to the lateral binning (summing up intensity between a predefined set
314 of borders) of 2D data into digital image elements and the step size refers to the raster of the sampling
315 stage or beam deflections⁵⁷. Step sizes smaller than the spot size were found to generate lower quality
316 images when sampling with a laser which does not ablate all ionizable compounds from one spot. In
317 this case, the tissue area is sampled with high overlap in adjacent sampling positions and sampling
318 from the subsequent spot will result in some signal from compounds of the previous spot position.
319 This is called oversampling. When the sampling area is completely ablated at each position without
320 oversampling, the overlapping position will not be cross contaminated and leads to a clear image. In
321 this case the resolution of the image is determined by the step size, since for each spot the sampled
322 ions originate from the non-overlapping and non-ablated sample area. For this situation the lateral
323 resolution is not limited by the diameter of the probe beam, but the intensity of the sample compound
324 will be lower due to the lower amount of available material in the non-ablated sample area^{116,117}.

325 **1.3.2 Pre-processing and visualization of large MSI data**

326 The data pre-processing, visualisation and interpretation depends on the dimensionality of the MSI
327 data. Tissue specimen has 3 dimensions (3D), from which a planar 2-dimensional (2D) tissue section
328 with defined thickness (generally 5-10 μm) is used for MSI. Orientation of the tissue section used for
329 MSI should be provided by sampling using an anatomical orientation description¹¹⁸. If multiple
330 adjacent tissue sections are analyzed then volumetric MSI data is acquired¹¹⁹. The dimensionality of
331 the MSI data is generally reflected as the spatial dimensions of the analysed tissue, thus it can be 2D
332 or 3D. MSI data obtained from a single tissue section is multidimensional with two spatial, one
333 separation (m/z) dimension and one quantitative readout (ion intensity). The two spatial dimensions
334 are in the plane of the analysed tissue section and the separation dimension consist of the mass-to-
335 charge (m/z) separation. The ion intensity is the quantitative readout, which is used for quantification
336 of the measured compounds.

337 “Pseudo” MSI data can be obtained by taking individual samples at different parts on an organ, the
338 full body or body surface and analyzing the samples with LC-MS or MALDI time-of-flight (TOF).
339 Mapping the measured data to the original sample location enables coarse mapping of compound

340 distribution in the analyzed samples, as shown in a study measuring metabolites, peptides and proteins
341 in samples taken from skins of volunteers by Bouslimani *et al*¹²⁰.

342 The size of the MSI data collected on large tissue sections at high spatial resolution is large and ranges
343 typically from 1 to 100 GB and in extreme cases can reach 1 TB such as for 3D FTICR data, but
344 smaller data sets of a few to hundreds of MB targeting small tissue areas with low spatial resolution
345 is collected routinely. There are many ways to pre-process, analyze and evaluate the large amount of
346 MSI data, and the main aims are to obtain a better understanding of the underlying molecular
347 mechanisms of biological events such as: (1) to determine the spatial distribution of compounds and
348 how this correlates with the anatomic morphology and cellular composition of the tissue, (2) to
349 determine how the spatial distribution of a particular compound correlates with those of the other
350 compounds. In the data interpretation process, visualization plays an important role, which is
351 challenging for the large amount of MSI data, but large data sets represent a challenge for pre-
352 processing as well. In order to reduce the volume of data, many data pre-processing pipelines use data
353 reduction techniques such as centroiding, noise reduction, intensity filtering and baseline removal,
354 creating images for features (isotopes) detected in a minimum number of spectra or filtering out ion
355 images that have low information content^{121,122}.

356 Suits *et al.*^{123,124} presented an approach which does not use any data reduction and allows to process
357 large volume profile MSI dataset as it was collected. This is achieved by using three different types
358 of indexed data structures of the same MSI data to allow interactive data interpretation by the users
359 without loss of information (**Figure 3**): (1) one representation contains sliced MSI data in the m/z
360 dimension with user defined thickness for fast visualization of MSI ion images, and enables
361 correlation queries between slices to find compounds that show spatial correlation with each other or
362 with an anatomical location, (2) representation of all MSI data in triplets of m/z , intensity and pixel
363 index. In this data, triplets are sorted and indexed according to m/z values, which serves to recalculate
364 a slice in the m/z dimension quickly with user defined thickness and m/z limits using a graphical user
365 interface, and (3) indexed storage of all MS spectra of each image pixel serving to quickly get MS
366 spectra for a particular tissue location.

367 The next level of data analysis is based on clustering similar mass spectra to determine how the spatial
368 distribution of the mass spectra clusters correlates to anatomic structures, a process called
369 segmentation. Another bioinformatics task is the alignment of the histology image with MSI data,
370 which transfers identified anatomical regions in the histology image to the MSI data and enables
371 identification of compounds in the identified anatomical region. This procedure is called the image
372 registration process and performs 2D alignment of the histology image to e.g. a specific m/z slice or
373 to the total ion current image (the sum of all ion intensities collected for one pixel) of MSI data¹²⁵.

374 Visualization of 2D MSI data obtained from one tissue slice is already challenging since one pixel is
375 described with four values (x and y coordinates, m/z value and intensity) and the most common
376 approach is to provide 2D images of single (normal image showing intensity as a color map) or
377 multiple (separate red, blue and green color maps combined with intensity dependent transparency
378 for 3 different slices) m/z slices. Visualization of ion intensity for a particular m/z range in 3D MSI
379 data obtained from a volumetric sample or visualization of multiple m/z slices in 2D MSI can be
380 performed with volumetric rendering frequently used in 3D computer graphics. Volumetric rendering
381 is a 3D visualization method for 4D data where the color and the transparency of one pixel is set
382 according to its intensity values (pixels with lower intensity are more transparent than pixels with
383 higher intensity)¹⁰⁹.

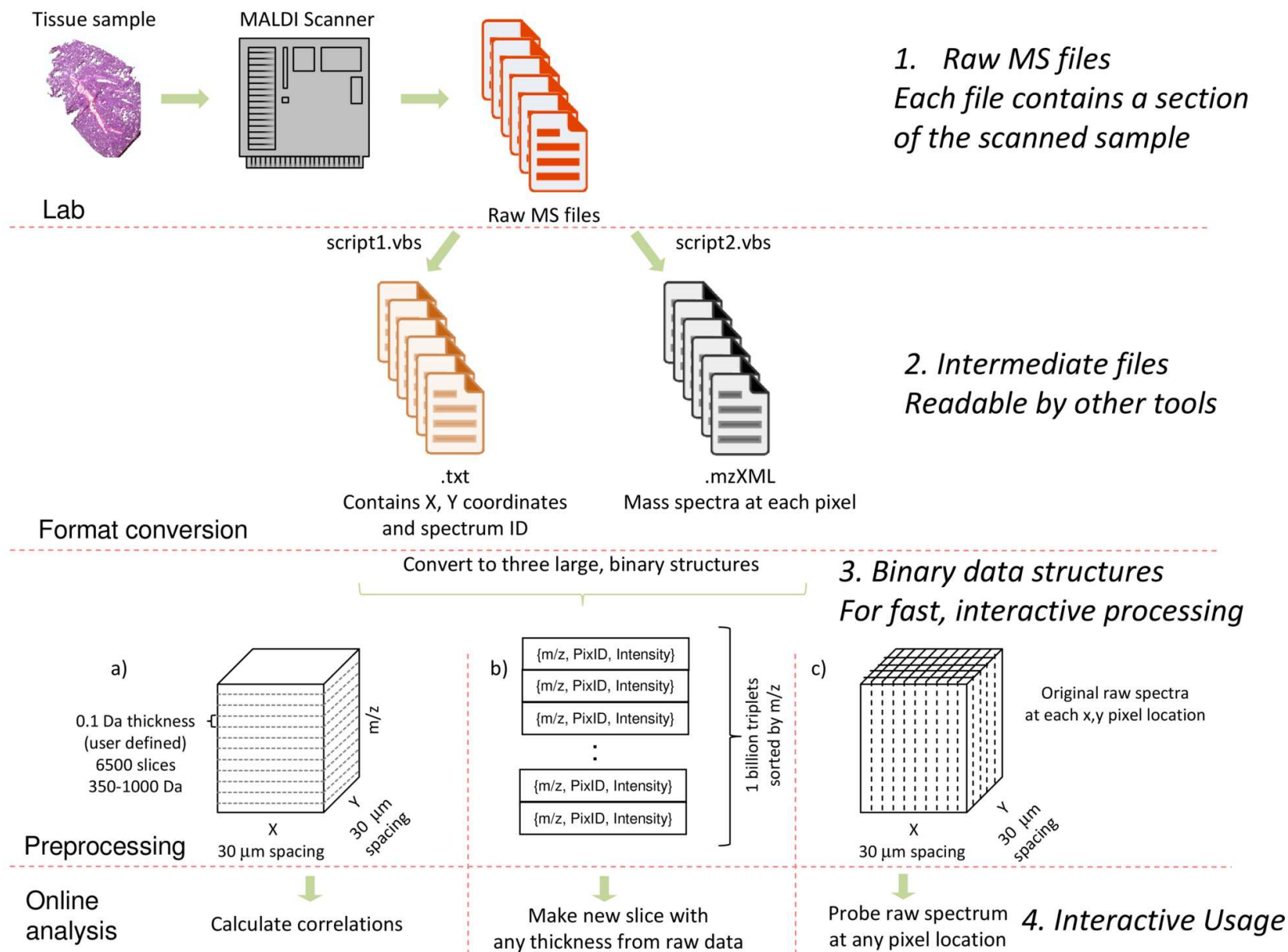


Figure 3. Data processing workflow, which allows analysing of all signals collected in an MSI experiment using an Orbitrap Velos instrument equipped with a MALDI interface interactively without information loss. Reprinted with permission from Suits *et al*¹²³. Copyright (2013) American Chemical Society.

387 Identification of the detected peptides and proteins by MSI is still challenging, for instance due to the
388 presence of isobaric compounds, poor fragmentation of large proteins, the presence of metabolites,
389 adduct formation and the presence of non-tryptic peaks when local trypsin digestion is performed on
390 the tissue section. Improvement of mass spectrometer sensitivity will allow detection of lower
391 abundant proteins, but may actually exacerbate the identification challenge by increasing spectrum
392 complexity. With a tandem mass analyzer, ions of interest can be specifically targeted for
393 fragmentation, to facilitate their identification. However, tandem mass spectrometry (MS/MS)
394 spectra of sufficient quality can only be obtained for ions with high intensity signals. An open non-
395 reviewed database, the MSiMass list (<https://ms-imaging.org/wp/msi-mass-list>), helps users to assign
396 identities to peaks submitted to MS/MS fragmentation observed in MALDI MSI experiments. This
397 database is the result of a community effort without a formal review panel and therefore information
398 in this database should be considered with care. In this concept, authors can freely enter data and can
399 comment on existing entries. Its ability to provide high quality data and identification is currently
400 under evaluation¹²⁶. In this section we have mentioned only the most important aspects and challenges
401 of MSI data processing and the reader is referred to a recent detailed review by Alexandrov on this
402 topic¹²⁷.

403 MSI data is acquired with a wide range of MS systems and many software tools are available to
404 process MSI data. For MSI data processing, imzML^{128,129} is the accepted open standard format, which
405 is supported by the Proteomics Standards Initiative of the Human Proteome Organisation (HUPO-
406 PSI)¹³⁰, and has become widely used for the flexible exchange and processing of MSI data between
407 different instruments and data analysis software. High-resolution molecular profiles of tissue
408 collected from MSI experiments often have data files of sizes of several tens to hundreds of gigabytes
409 requiring powerful visualization software, such as the Biomap (Novartis, Basel, Switzerland,
410 www.maldi-msi.org) image processing application, the MALDI Imaging Team Imaging Computing
411 System (MITICS)¹³¹ and the Datacube Explorer (DCE, available at the www.imzml.org) to explore
412 imaging mass spectrometry data sets¹³². Recently, high-quality 3D MALDI and DESI benchmark
413 MSI datasets in imzML format were made available for software evaluation purposes¹³³.

414 **2 Untargeted mass spectrometry imaging of proteins**

415 This section discusses MSI strategies for hypothesis-free untargeted analysis of protein distribution
416 in tissue and presents the technological limitations and current challenges, illustrated with example
417 applications. Untargeted analysis of protein distribution requires the collection of ion intensity signals
418 specific to proteins and linking accurate identification to these signals. In untargeted MSI, proteins
419 can be identified with two approaches. In the first approach the proteins are digested *in situ* by

420 application of a protease (typically trypsin) in isolated spots, and the proteins in these spots are
421 cleaved into peptides. These peptides are then ionized, sampled into the mass spectrometer and
422 analyzed intact or following fragmentation using conventional MS/MS fragmentation methods such
423 as collision induced (CID) or electron transfer dissociation (ETD). The application of droplets limits
424 the spatial resolution of this approach. The second approach uses ionize, sample into the mass
425 spectrometer and analyze intact proteoforms, which can be combined with fragmentation methods
426 such as higher-energy collision induced dissociation (HCD) and ETD that can be directly applied to
427 intact proteins extracted from tissue and submitted to purification^{134–136}. The first strategy does not
428 provide information on the entire protein sequence, and the detected peptides in many cases do not
429 allow differentiation between protein isoforms or partially degraded proteins in the absence of
430 additional information (e.g. the mass of the intact protein). Top-down fragmentation of intact protein
431 provides more complete information on the entire protein sequence and allows better discrimination
432 between isoforms, but requires clean and extracted proteins and cannot be applied directly in an MSI
433 setting. The advantage of the first approach is that it can be applied to determine the distribution of
434 post-translational modifications of specific residues in proteins directly from tissue¹³⁷.

435 **2.1 Untargeted MALDI MSI of intact proteins in tissue**

436 **2.1.1 Extending the mass range for intact protein MALDI MSI**

437 The matrix deposition method has a critical impact on the mass range of intact protein MSI.
438 Leinweber *et al.* developed a sandwich matrix deposition protocol, which includes application of
439 different solvents and detergents for MALDI MSI of proteins in tissue sections, extending the mass
440 range to 25-50 kDa and increasing the number of detected intact proteins. This protocol uses two
441 layers of matrix, one below and one on top of the tissue section, and has been employed for MSI of
442 proteins in kidney, heart, lung and brain tissue sections of different rodent species¹³⁸. Grey *et al.*
443 introduced a tissue preparation procedure, which includes extensive washing with water to remove
444 highly abundant water-soluble proteins, and automated spotting of matrix solution using a high
445 percentage of organic (acetonitrile) solvent. This protocol allowed to measure membrane proteins up
446 to 28 kDa in bovine lens, human lens, and rabbit retina by MALDI MSI, but at moderate spatial
447 resolution of 100-200 μm due to application of matrix spotting¹³⁹. Franck *et al.* enhanced the
448 solubilization of large proteins using hexafluoroisopropanol (1,1,1,3,3,3-hexafluoro-2-propanol) and
449 2,2,2-trifluoroethanol during sample preparation and achieved MSI of proteins between 30 and 70
450 kDa directly from tissue¹⁴⁰. Mainini *et al.* investigated ferulic acid as matrix on different tissues
451 deposited with an automated matrix deposition device, ImagePrep (Bruker Daltonics, Bremen,
452 Germany), which performs matrix deposition by spraying sequences and allowed the detection of
453 proteins up to 135 kDa¹⁴¹.

454 The shortcoming of widely-used mass spectrometers is the inefficient transmission and fragmentation
455 of large proteins^{138,140–144}, particularly the low transmission efficiency of the latter. Recent
456 development of mass spectrometers has enabled the implementation of large protein analysis even
457 under native conditions by enhancing the ion transmission of intact proteins up to one megadalton¹⁴⁵.
458 These developments have allowed to expand the mass range within which intact proteins can be
459 analyzed and will certainly contribute to generate more informative MSI data.

460 Another improvement of MSI of intact proteins was achieved by van Remoortere *et al.*, who used a
461 high mass HM1 TOF detector (CovalX, Zurich, Switzerland) to improve the sensitivity of MALDI
462 MSI of intact proteins up to 70 kDa¹⁴⁶. Compared with traditional micro-channel plate detectors, this
463 instrument has a much larger charge capacity and is therefore less prone to detector saturation.
464 Another novel method in MALDI MSI was described by Jungmann *et al.*, who used a parallel, active
465 pixel TOF detector for MSI of ubiquitin oligomers reaching a molecular mass of 78 kDa¹⁴⁷.

466 Although these methods demonstrate encouraging results for imaging proteins of increasing mass,
467 each of these protocols has some drawbacks that are usually associated with low reproducibility,
468 including: ion suppression effects¹⁴⁸, low ion yield (it has been estimated that only 1 molecule ionizes
469 out of 1000 desorbed proteins^{149–151}), the need for a special non-commercially available mass
470 analyzer¹⁴⁶, a limitation to detect membrane proteins¹³⁹, the requirement of complex and laborious
471 experimental protocols¹³⁸ and time-consuming, as well as extensive sample preparation¹⁴⁰.

472 **2.1.2 Spatial resolution improvement of MALDI MSI for intact proteins**

473 A number of methods were developed to implement the spatial resolution of MALDI MSI of proteins
474 from tissue samples. These methods focused on improving the sample preparation protocol, reducing
475 the laser beam spot size, and improving the ion sampling and transmission parts of the mass
476 spectrometer. As mentioned in section 2.1.1, tissue sample preparation is the most important factor
477 to achieve both high sensitivity and high spatial resolution in MALDI MSI. McDonnell *et al.*
478 performed an extensive comparison of five tissue washing protocols using human arterial tissue
479 samples, and assessed the methods in terms of the information content (e.g. number of detected peaks,
480 quality of morphological structures) as well as their suitability for analyzing tissue containing small
481 but distinct regions. In this work, they demonstrated an optimized tissue washing protocol using 70%
482 and 90% isopropanol for imaging proteins that are specific to the *intimas* and *media* layers of
483 atherosclerotic arterial tissues at a high spatial resolution of 30 μm ¹⁵². With an appropriate laser spot
484 profile (flat-top) and diameter (10-20 μm) and a matrix application method (spraying matrix with the
485 Bruker ImagePrep device) that precludes analyte delocalization and maintains the original lateral
486 spatial distribution of proteins, the group of Pineau reported a MALDI MSI of proteins in the 10 kDa
487 range in rat testis tissue at 20 μm lateral resolution¹³⁶. Caprioli's group implemented a matrix

488 sublimation/recrystallization process, which provides a more homogeneous distribution of the matrix
489 resulting in more sensitive detection of large proteins using MALDI MSI with a spatial resolution as
490 low as 10 μm ⁸⁴. Additionally, for targeted analysis, histology-directed imaging was performed using
491 this protocol, where MSI analysis and hematoxylin and eosin (H&E) staining were performed on the
492 same tissue section which was previously used for MSI. Integrating H&E staining with MSI data
493 acquired on the same tissue section allows to transfer anatomical annotation from H&E staining to
494 MSI data and allows to identify protein signals which correlate spatially with anatomical features. In
495 another study, Deutskens *et al.* applied a robotic spray apparatus for matrix application, and
496 performed MALDI MSI on a tissue section followed by elimination of the matrix by washing and
497 subsequent histology staining and microscopic examination of the same tissue section. This matrix
498 application protocol has two steps (one dry matrix coating and one hydration/recrystallisation), which
499 separates the processes of matrix coating from analyte extraction and provides a highly reproducible
500 homogenous matrix layer. A key advantage of this protocol is that it limits the delocalization of
501 proteins and enables imaging at a relatively high spatial resolution of 35 μm ¹⁵³.

502 The spatial resolution achievable with MALDI is ultimately restricted by the size of the laser spot¹⁵⁴.
503 While it is possible to image with a spatial resolution less than the diameter of the laser beam by
504 oversampling (i.e. with a laser spot size of 60 μm , one could raster with 20 μm steps) to effectively
505 achieve 20 μm spatial resolution¹¹⁶, it is important to completely ablate the prior spot before moving
506 the laser beam to the next position to reduce crosstalk between pixels. To minimize the laser spot size,
507 the group of Caprioli *et al.* developed a new source for MSI with a transmission geometry that allows
508 the laser beam to irradiate the backside of the sample and the separation of ion and laser optics
509 resulting in a laser spot size close to the wavelength of the applied laser, thereby allowing MSI at
510 higher spatial resolution. This method produced high-quality images of intact insulin in the cytoplasm
511 at sub-cellular resolution in mouse cerebellum tissue¹⁵⁵. With appropriate sample preparation and
512 using 2,5-dihydroxyacetophenone as matrix, the transmission geometry principle was able to achieve
513 a 1 μm laser spot diameter on target with a minimal raster step size of 2.5 μm . This approach allowed
514 to produce mass spectrometry images of proteins acquired in a step raster mode at 5 pixels/s and in a
515 continuous raster mode at 40 pixels/s¹⁵⁶, which is much faster than the 0.5-2 pixel/s acquisition of
516 common QTOF and Orbitrap instruments. Increasing acquisition speed has the advantage that data is
517 acquired within a reasonable time frame, which prevents molecular alteration of tissue in time from
518 the beginning to the end of the MSI process. Zavalin *et al.*¹⁵⁷ developed a “laser beam filtration”
519 approach, using lenses and a 25 μm ceramic spatial filter (pinhole) to remove the satellite secondary
520 laser beam energy maxima resulting in a well-defined 5 μm diameter laser spot. The images generated
521 from a mouse cerebellum showed clearly distinguishable cellular forms such as the Purkinje layer,
522 dendrites, and axon fibers. Spengler’s group introduced a Scanning Microprobe Matrix Assisted

523 Laser Desorption/Ionization (SMALDI)-MSI method, which features the possibility to investigate
 524 and visualize the spatial distribution of analytes including peptides such as bradykinin and angiotensin
 525 II in samples with sub-cellular resolution (0.5-10 μm) in pine tree roots^{107,158}.

526 Spatial resolution MSI of proteins from tissue sections can also be improved with specific sample
 527 preparation techniques or with dedicated data processing. Caprioli *et al.* have developed an approach
 528 to image proteins by blotting the tissue sections on a specially prepared target containing an adsorbent
 529 material⁸⁰. Peptides and small proteins bind to the C₁₈ material and create a positive imprint of the
 530 tissue, which can then be imaged by the mass spectrometer. The imprinted tissue material prevents
 531 any further delocalization of proteins and enables washing away interfering compounds such as lipids
 532 and salts. This approach has been applied to map proteins from the rat pituitary gland with a spatial
 533 resolution of $\sim 25 \mu\text{m}$. Integration of a coaxial laser illumination ion source into a MALDI-TOF-MS
 534 instrument allowed visualization of proteins of a molecular mass up to 27 kDa using this approach.
 535 In another study, two highly expressed secretory epididymal proteins in a mouse caudal epididymis
 536 tissue section were visualized, with a spatial resolution below $10 \mu\text{m}$ ⁹².

537 Low spatial resolution MSI data can be combined with high-resolution spatial microscopic images
 538 using multivariate regression called image fusion approach. Image fusion enables to predict
 539 distribution of MSI data at the spatial resolution of the H&E image. The resulting images combine
 540 the advantages of both technologies, enabling prediction of a molecular distribution both at high
 541 spatial resolution and maintaining the high chemical specificity of MSI data. For example, an ion
 542 image of m/z 778.5 (identified as a lipid) measured in mouse brain at $100 \mu\text{m}$ spatial resolution, can
 543 be extrapolated for $10 \mu\text{m}$ spatial resolution using fusion with H&E microscopy image measured
 544 from the same tissue sample at $10 \mu\text{m}$ resolution. Another example describes the prediction accuracy
 545 of an ion image with m/z 10,516 Da corresponding to an unidentified protein measured in a mouse
 546 brain section at $100 \mu\text{m}$ resolution and fused with an H&E microscopic image resulting in a predicted
 547 image at $5 \mu\text{m}$ resolution. This approach has been successfully applied for various tissue types, target
 548 molecules and histological staining protocols at different resolution scales. In addition, this approach
 549 can generate ion image predictions using microscopic images at the nanometer range, below the
 550 resolution achievable with current MALDI MSI instrumentation¹⁵⁹. However, it should be noted that
 551 the image fusion approach is a statistical procedure predicting distribution at higher spatial resolution
 552 than the actually acquired MSI data. Therefore, thorough assessment of the prediction accuracy
 553 should be applied for each specific location and m/z slice.

554 A study from Spraggings *et al.*¹⁶⁰ presents an ultra-high speed MALDI-TOF MS, which provides
 555 image acquisition rates >25 pixels/s with high spatial resolution of 30 (full tissue section) and $10 \mu\text{m}$
 556 (only selected tissue areas due to time required to collect the data) and a high mass resolution MALDI

557 Fourier transform ion cyclotron resonance (FTICR) MS operated with 100 μm spatial resolution.
558 These novel instruments improve protein image acquisition rates by a factor of 10, can provide
559 MALDI MSI data at 10 μm spatial resolution with good sensitivity, and isotopically resolve proteins
560 up to 20 kDa. The data from these two instruments on the same tissue section could be combined e.g.
561 with interpolation similar to the image fusion approaches resulting in high spatial resolution and high
562 mass accuracy MSI data.

563 **2.1.3 Identification of intact proteins in MSI**

564 Intact proteins can be fragmented in the gas phase outside or inside the mass spectrometer through
565 various mechanisms¹⁶¹, such as MALDI in-source decay (ISD), collision-induced dissociation (CID),
566 infrared multiphoton dissociation (IRMPD), electron capture dissociation (ECD), ETD, ultraviolet
567 photodissociation (UVPD) and laser-induced dissociation (LID). Among these, MALDI ISD^{162–164},
568 where the fragmentation occurs in the MALDI ion source is the most widely used approach^{80,76,165}.
569 ISD has proven to be an efficient method for the N- and C-terminal sequencing of proteins in tissue
570 sections. In ISD, proteins are cleaved at the N-C α bond of the peptide backbone at high laser fluence
571 (radiant exposure expressing the amount of energy received per unit of surface area) in the hot
572 MALDI plume, giving principally c- and z-type protein fragments¹⁶⁶. As early as 2001, Chaurand *et al.*
573 applied ISD-MSI in the characterization of spermine-binding protein (SBP) in mouse prostate
574 lobes with respect to sequence variants and PTMs and the localization of this protein¹⁶⁷. The main
575 advantage of ISD is that there is no mass limitation since fragmentation occurs prior to ion
576 acceleration. However, ISD suffers from the major drawback of lack of precursor ion selection, which
577 leads to a complicated mass spectrum if more than one protein is present at the laser shot position,
578 which is generally the case in MSI of tissue section. In addition, many c- or z-fragment ions below
579 1000 Da are often difficult to assign due to the presence of matrix adduct peaks, making the
580 identification of the sequence part close to the protein termini challenging. ISD-MSI require multiple
581 laser shots in the same spot ablating all available proteins to gain the highest signal, which is a time-
582 consuming task.

583 To circumvent this issue, a “pseudo-MS³” approach, also known as “T³-sequencing”, has been
584 developed to improve MALDI-ISD in proteins^{168,169}. In this approach, the fragments produced by
585 ISD are further isolated and fragmented with a classical tandem MS/MS approach in QTOF or
586 MALDI-TOF/TOF instruments. The T³-sequencing method with specific MALDI matrices, such as
587 2,5-dihydroxybenzoic acid or 1,5-diaminonaphthalene, has been applied to identify proteins such as
588 myelin basic protein and crystallins in the tissue slices of mouse brain and porcine eye lens
589 respectively¹⁶³. The efficacy of MALDI-ISD-MSI to simultaneously identify the protein and
590 determine its localization has been demonstrated in another study using tissue sections of porcine eye

591 lens. In this study a new bioinformatics pipeline was presented for processing MALDI-ISD-MSI data
592 to identify proteins based on spectra containing high numbers of correlated fragments that are likely
593 to be part of the same protein. This approach allows to determine the lateral spatial distribution of
594 identified proteins as well¹⁷⁰. Pauw and coworkers recently presented a high-resolution MALDI-ISD-
595 FTICR method to identify a set of selected protein markers on histological slices simultaneously with
596 minimal sample pretreatment¹⁷¹. In this method, known protein markers are spotted next to the tissue
597 of interest and the whole MALDI plate is coated with 1,5-diaminonaphthalene matrix. The latter
598 promotes MALDI ISD, providing large amino acid sequence tags. Comparative analysis of ISD
599 fragments between the reference spots and the specimen in imaging mode allows for unambiguous
600 identification of protein markers while preserving full spatial resolution, as well as the N- and C-
601 terminal sequencing of proteins present in tissue sections. This was demonstrated with the distribution
602 of myelin basic protein (MBP) from mouse brain and human neutrophil peptide 1 (HNP-1) in human
603 liver sections containing metastasis from colorectal cancer.

604 Another approach to identify proteins uses fragmentation methods in mass spectrometers applied in
605 “top-down” protein analysis such as ETD, ECD, or UVPD^{172–175}. These might be applicable to top-
606 down identification approaches in MSI, although the speed and sensitivity are currently not yet
607 compatible with MSI. Even with these novel achievements, the detection of signals from intact
608 proteins will still remain much easier than performing accurate identification, which will result in the
609 fact that the majority of the protein signals in MSI remains unidentified.

610 **2.2 Mass spectrometry imaging of proteins after *in situ* digestion**

611 Another strategy used in MSI for protein imaging is *in situ* digestion prior to MALDI MSI analysis,
612 which can be used to identify proteins and to determine protein distribution using surrogate
613 proteotypic peptides. The method retrieves protein distributions in tissue sections using the
614 corresponding proteotypic peptides after enzymatic digestion, most of the time using trypsin.
615 Proteotypic peptides are those peptides that uniquely identify a protein and are used in bottom-up
616 targeted and untargeted proteomics workflows to identify and quantify proteins with a (tandem) mass
617 spectrometer¹⁷⁶. In fact, peptides are smaller and, due to their better fragmentation, are easier to be
618 identified by tandem mass spectrometry. Additionally, peptide fragments are easier to obtain than
619 intact proteins from FFPE tissue. Therefore *in situ* digestion analysis is the method of choice for this
620 sample type, which is more abundantly available in hospital biobanks compared to fresh frozen tissue
621 samples. With this technique, Caprioli and coworkers described on-tissue identification of proteins
622 in spatially discrete regions using tryptic digestion followed by MALDI MSI with (TOF-TOF)
623 MS/MS analysis¹⁷⁷. The procedure in this study identified several proteins in the coronal sections of
624 a rat brain including higher molecular weight proteins, such as actin (41 kDa), tubulin (55 kDa), and

625 synapsin-1 (74 kDa). Ronci and Voelcker applied on-tissue trypsin digestion to analyze the freshly
 626 excised human lens capsule by MALDI MSI. This work demonstrated that the distribution of proteins
 627 can be determined from this highly compact connective tissue having no evident histo-morphological
 628 characteristics. Furthermore, the study shows a high repeatability of the digestion protocol on four
 629 different human lens capsule specimens by evaluating the distribution of the same set of peptides¹⁷⁸.
 630 Recently, Diehl *et al.* optimized the *in situ* imaging of protein distribution after protease digestion
 631 with MALDI MSI using cryoconserved and FFPE rat brain tissue by applying different digestion
 632 times, types of matrix, and proteases¹⁷⁹. The conclusion of this study was that the digestion time does
 633 not play an important role for the quality of MSI images, while trypsin provided the highest number
 634 of peptide signals corresponding to anatomical regions.

635 Ion mobility separation (IMS) combined with MSI has emerged as a powerful technique to improve
 636 specific detection of isobaric peptides with different molecular shape^{25,180–182}. For example, Clench
 637 and coworkers successfully performed IMS-MSI to localize and identify peptides of the glucose-
 638 regulated protein 78 kDa (Grp78), which is known as a tumor biomarker, directly from FFPE
 639 pancreatic tumor tissue sections. Grp78 was found to be mainly located in tumor regions using
 640 MALDI-IMS-MSI¹⁸¹. In this procedure IMS separated isobaric peptides, which facilitated their
 641 identification following fragmentation, obtaining a cleaner image with less interferences for a
 642 particular peptide. Stauber *et al.* applied enzymatic digestion protocols for MALDI-IMS-MSI with
 643 high sensitivity localization and identification of proteins from FFPE and frozen tissues obtained from
 644 rat brain¹⁸². This study showed that isobaric peptides can be separated, which improves ion image
 645 specificity and improves identification accuracy of fragmented peptides.

646 Schober *et al.* presented a method for imaging tryptic peptides¹⁸³ in which MALDI MSI experiments
 647 were complemented by off-line liquid chromatography coupled to electrospray ionization tandem
 648 mass spectrometry (LC-ESI-MS/MS) analysis on an FT-ICR mass spectrometer to increase the
 649 number of identified peptides and proteins. Comparative results were obtained by analyzing two
 650 adjacent mouse brain sections in parallel. The first section was spotted with trypsin and analyzed by
 651 MALDI MSI. On-tissue MS/MS experiments of this section resulted in the identification of only 14
 652 peptides (originating from 4 proteins). The second tissue section was homogenized, fractionated by
 653 ultracentrifugation and digested with trypsin prior to LC-ESI-MS/MS analysis. The number of
 654 identified peptides increased to 153 (corresponding to 106 proteins) by matching imaged mass peaks
 655 to peptides which were identified in these LC/ESI-MS/MS experiments. This identification difference
 656 can be explained that selected precursor ion windows in direct fragmentation of peptides from tissue
 657 include matrix and other interference which results in noisier spectra compared to LC-MS/MS
 658 analysis where these interferences are not present.

659 The group of McDonnell reported a comprehensive study of the mouse brain proteome from mouse
660 brain slices with MSI using multiple proteases such as trypsin, Lys-C, Lys-N, Arg-C, and a mixture
661 of trypsin and Lys-C¹⁸⁴. This study combined identification of peptides and proteins from tissue using
662 bottom-up LC-ESI-MS/MS and linked the obtained identifications using accurate mass with non-
663 fragmented MSI data. In the LC-ESI-MS/MS data 5337 peptides were identified using
664 complementary proteases, corresponding to 1198 proteins. 630 of these peptides, corresponding to
665 280 proteins, could be assigned to peaks in MSI data sets and used to determine the parent protein
666 distribution in tissue. Gene ontology and pathway analyses revealed that many of the proteins are
667 involved in neuro-degenerative disorders, such as Alzheimer's, Parkinson's, and Huntington's
668 disease¹⁸⁴, which highlights the potential application of the technique in the future for diagnosis and
669 pathology purposes.

670 Many approaches have been developed to improve protein identification performance in MALDI
671 MSI after enzymatic digestion. For example, Franck *et al.* developed an N-terminal chemical
672 derivatization strategy using 4-sulphophenyl isothiocyanate (4-SPITC), 3-sulfobenzoic acid (3-SBA)
673 and N-succinimidylloxycarbonylmethyl-tris(2,4,6-trimethoxyphenyl)phosphonium bromide (TMPP)
674 reagents, which improves *de novo* peptide identification performance¹⁸⁵. The reagents added an
675 additional positive or negative charge at the N-terminus of tryptic peptides, which provided more
676 complete ion series upon fragmentation. From these reagents TMPP provided the best performance
677 in terms of fragmentation efficiency of peptides from tissue. Clench's group used a recombinant
678 protein termed "IMS-TAG" for MALDI-IMS-MSI²⁵. The IMS-TAG recombinant proteins are
679 engineered and used as a multi-protein standard. After trypsin digestion, this IMS-TAG protein yields
680 – analogous to the QconCAT¹⁸⁶ approach – a range of peptides that can be used as internal standards
681 to identify and quantify multiple proteins in a MALDI-IMS-MSI experiment. In this approach IMS
682 is used to provide an additional selectivity to detect IMS-TAG derived standard peptides and to
683 remove any potential interfering isobaric peptide signals. In this study, MALDI-IMS-MSI was used
684 to measure the distribution of HSP90 and vimentin in FFPE EMT6 mouse tumor sections, as well as
685 HSP90 and plectin in a fresh frozen mouse fibrosarcoma using extracted ion images at the
686 corresponding *m/z* values and drift times from IMS-MSI data.

687 Performing accurate protein quantification in MSI is challenging since ion suppression due to other
688 co-localized compounds can be strong and protein extraction and desorption can be partial in case of
689 MSI of intact proteins. Trypsin digestion may alter quantification since this step creates a new ion
690 suppression environment. The quantification performance can be made more accurate by using spiked
691 stable isotope standards. For example, Porta *et al.* used stable isotope standards and performed
692 quantification based on fragment ions obtained in SRM mode, which allowed to achieve a

quantification precision of 10-15%, which is sufficient to meet requirements of most bioanalysis guidelines¹⁸⁷. A further finding of this work was that single pixel quantification is less accurate and at least the average of 4-5 pixels is required for accurate quantification of compounds in MSI data.

Komatsu *et al.* presented a feasibility study using a bismuth cluster ion (Bi_3^+) source with SIMS-TOF-MSI to determine protein distribution at the sub-cellular level combined with the ink-jet printing of trypsin. In this approach, a modified bubble jet printer (PIXUS 990i, Canon Inc.) was used to deposit trypsin and trifluoroacetic acid on a human serum albumin film layer. Protein images were obtained by visualizing the dot-patterned proteotypic peptide ions¹⁸⁸. Nygren and Malmberg mapped tryptic fragments of thyroglobulin (660 kDa) in pig thyroid glands after trypsin digestion by SIMS-MSI using a Bi_3^+ primary ion source. In this study, trifluoroacetic acid in water was used to improve the ionization of the peptides, which resulted in a 3 μm spatial resolution MSI image showing a heterogeneous distribution of this protein in the thyroid follicle cells⁴⁰.

3 Targeted mass spectrometry imaging of protein in tissue using tag-mass probes

This section presents approaches to circumvent some of the shortcomings of MALDI MSI of proteins and peptides by not using matrix and detecting proteins with targeted indirect signals resulting from chemical derivatisation and immunochemistry recognition. Two major approaches are discussed in this section: the use of LA-ICP for detecting metals in proteins and the Tag-Mass approach.

LA-ICP MSI generates signals for targeted biomolecular imaging, which can be applied for MSI of proteins with high sensitivity and dynamic range, but at a relatively low spatial resolution (100-200 μm). For example, Seuma *et al.* studied the distribution of two breast cancer-associated proteins, MUC-1 and HER2 in tissue sections by measuring Au or Ag tagged antibodies, but although successful it was concluded that the image quality was inferior to microscopy¹⁸⁹. Becker *et al.* demonstrated the potential of LA-ICP-MS to detect metalloproteins in protein bands or spots excised from 1D and 2D gel electropherograms. This method was then applied for sensitive and quantitative imaging of metals in brain sections, with detection limits for copper and zinc at the $\mu\text{g/g}$ tissue level and below¹⁹⁰. Giesen *et al.* applied LA-ICP-MSI for imaging metal-labelled antibodies to detect and quantify proteins directly in breast cancer and palatine tonsil tissue samples¹⁹¹. More recently, the same group developed this method further, and used 32 metal labeled antibodies to determine simultaneously 32 markers for protein and protein modification distribution in breast cancer tissue with laser ablation on a CyTOF instrument at subcellular resolution. The subcellular resolution at 1 μm enabled them to use this approach as mass spectrometry based cytometry i.e. to measure the concentration of these 32 protein markers in individual cells in tissue sections^{192,193}.

726 In 1998, a novel PC mass tag strategy for targeted detection of proteins has been suggested by Olejnik
727 *et al.*¹⁹⁴ This strategy implements the targeted analysis of proteins by affinity labeling with an antibody
728 (or another affinity agent) containing a PC mass tag and analyzing the labeled sample with LDI. The
729 tag contains a PC-linker, linking the antibody to the mass tag, which is cleaved upon LDI, released
730 into the gas phase, ionized and sampled into the mass spectrometer without the requirement to apply
731 matrix for the analysis. Due to the absence of matrix, spatial resolution is not limited by the size of
732 the analyte-matrix co-crystal and sensitivity is improved because detection of the released mass-tag
733 reporter fragment ion does not interfere with matrix cluster ions. In the absence of matrix, the spatial
734 distribution of LDI image is determined by the beam diameter of the applied laser. The PC-linker is
735 cleaved with high yield under the near-UV laser pulses commonly used in MALDI-MS instruments.
736 With a well-designed PC-linker and mass tag, this strategy has the ability to detect non-ionizing
737 compounds and offers high selectivity and sensitivity for target proteins. Furthermore, coupling
738 multiple PC-linked reporter mass tags to one affinity compound enhances the sensitivity of detection
739 by increasing the MS signal¹⁹⁵.

740 Although MALDI MSI has a much lower lateral resolution than classical optical microscopy ($\ll 1$
741 μm for example by using fluorescently labeled proteins), MS is both a sensitive method and allows
742 for the simultaneous (multiplexed) detection of hundreds to thousands of compounds. For
743 fluorescence, only a restricted number of fluorophores are available, whereas the number of mass
744 tags is only limited by the number of fragment ions that a mass analyzer can distinguish, which is *a*
745 *priori* almost unlimited. Therefore, the mass tag method is a promising matrix-free strategy, which
746 has a high multiplexing capacity, and the detection and localization of proteins in tissue sections with
747 high specificity and sensitivity allowing to detect proteins larger than 30 kDa. A limitation is the
748 availability of separate specific affinity reagents with unique mass tags for each protein to be
749 measured and the specificity of the affinity tag.

750 In the literature, two types of photolinkers and reporter fragments (mass tags) have been reported,
751 which have been developed by two different research teams. The group of Fournier described a
752 targeted PC-linker strategy termed Tag-Mass based on the photocleavable linker 4-[4-[1-(Fmoc-
753 amino)ethyl]-2-methoxy-5-nitrophenoxy]butanoic acid coupled to a peptide such as bradykinin as the
754 mass tag. To study the possibility of using photocleavage under multiplex analysis conditions, this
755 group used a mixture of three photocleavable-tagged oligonucleotide probes corresponding to three
756 different 20-mer oligonucleotides recognizing particular mRNA (**Figure 4**)²². Although 100%
757 photocleavage yield was not achieved using MALDI, the MS spectra showed the expected m/z of the
758 mass tag demonstrating efficient photocleavage by laser irradiation. To increase the sensitivity, the
759 group designed a new photocleavable linker/tag system by replacing the disulfide bridge with a

maleimide group for binding the peptide serving as mass tag to the photocleavable linker. This concept was applied to obtain specific images of proteins using tagged secondary antibodies. The results showed that MALDI appears to have a better sensitivity than the optical fluorescence images obtained from the same tissue section.

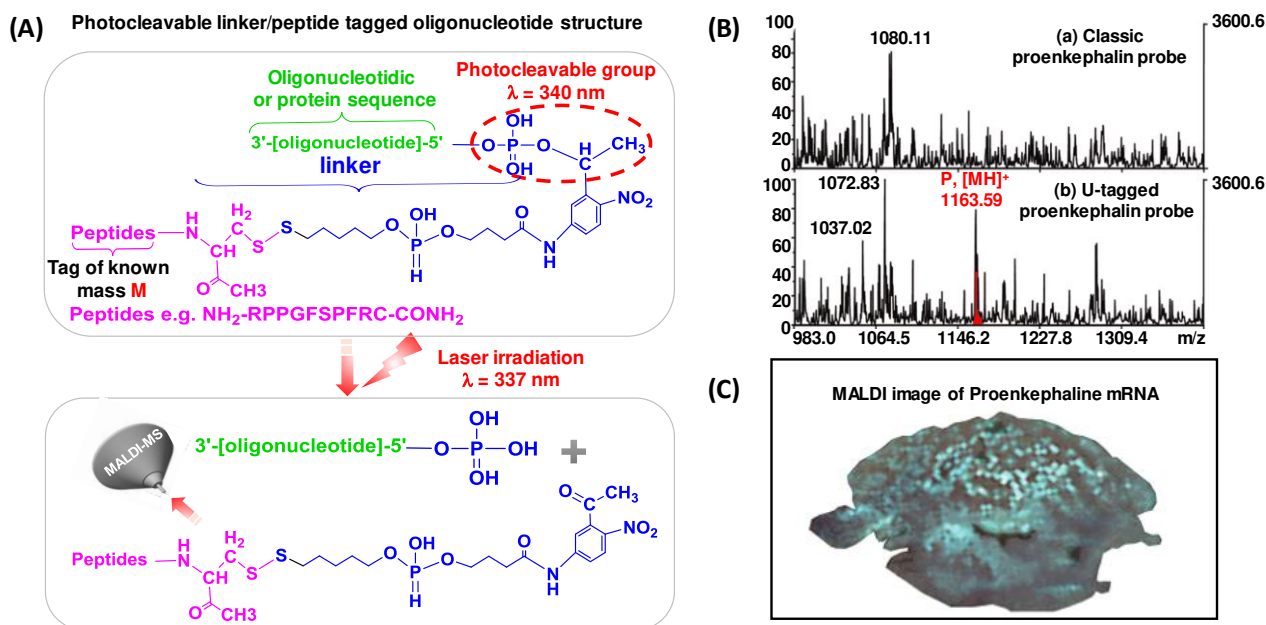
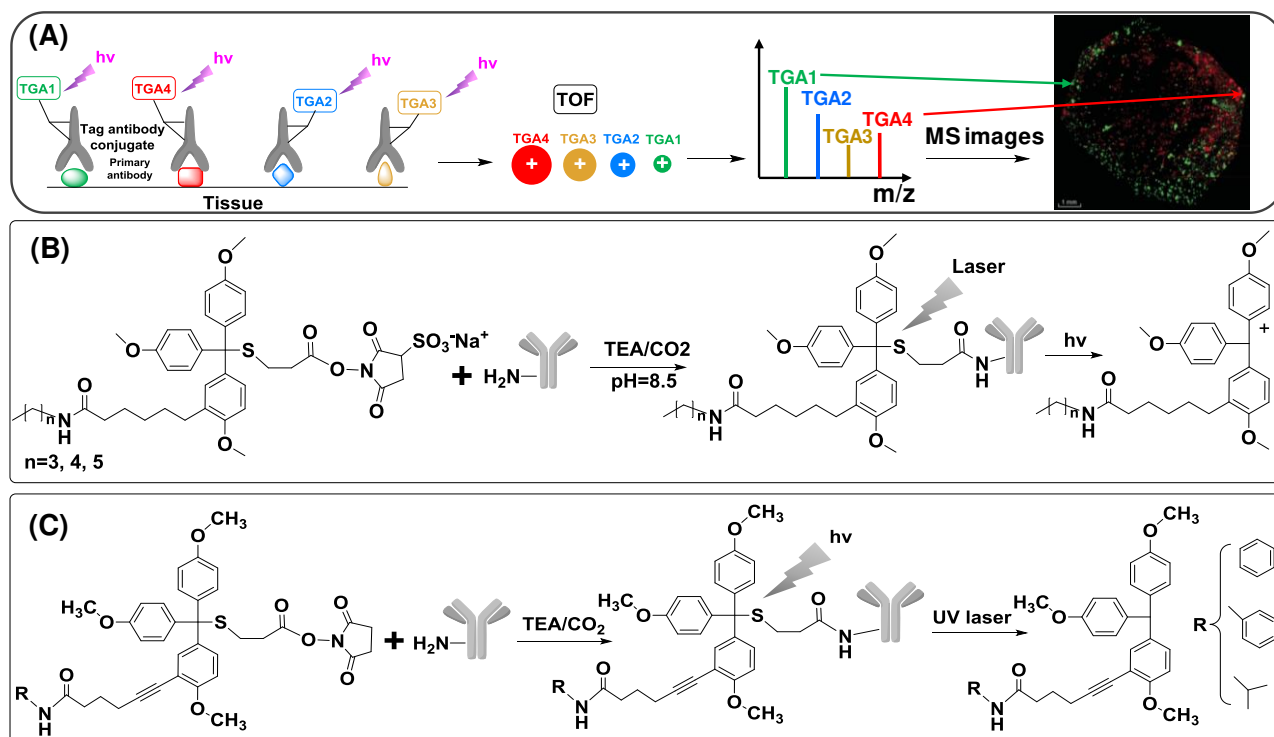


Figure 4. Structure of a photocleavable linker/tag system conjugated to an oligonucleotide/protein moiety and the reporter mass tag released via photocleavage as a result of irradiation by the UV laser (A). MALDI spectra of the untagged proenkephalin probe (upper plot) and the Uracil-tagged (U-tagged) proenkephalin probe (B) showing the peak highlighted in red corresponding to the applied mass tag in rat brain. Ion distribution image of the mass tag corresponding to the proenkephalin mRNA distribution. Adapted with permission from Lemaire *et al.*²². Copyright (2007) American Chemical Society.

The Tag-Mass strategy has been extended to different types of targeting compounds including secondary and primary antibodies, lectins and aptamers, which can be used to selectively obtain images of specific protein antigens, glycosylated proteins and drugs, respectively²⁴. It can be combined with hybridization and affinity recognition techniques including *in situ* hybridization of mRNA (ISH) and immunohistochemistry (IHC)^{22,24,196,197}.

In 2007, Thiery *et al.*¹⁹⁸ reported a novel photocleavable mass-tag approach, where the released tag can be detected under LDI conditions and used for TAMSIM (**Figure 5**). TAMSIM is based on an N-hydroxysuccinimide (NHS) linker coupled to trityl reporters with a thiopropionate group, which provides low molecular weight fragments (500-600 Da) in LDI¹⁹⁸⁻²⁰⁰. In this reagent, the trityl groups absorb UV light and form a resonance-stabilized carbocation, which results in cleavage of the C-S bond, and the release of the ionized mass-tag without the use of a matrix. This strategy was successfully applied to localize three different cancer markers on human tissue sections,

784 synaptophysin, protein S100 (PS100) and human melanosome (HMB45), that are normally below the
 785 detection threshold of untargeted MALDI MSI¹⁹⁸.



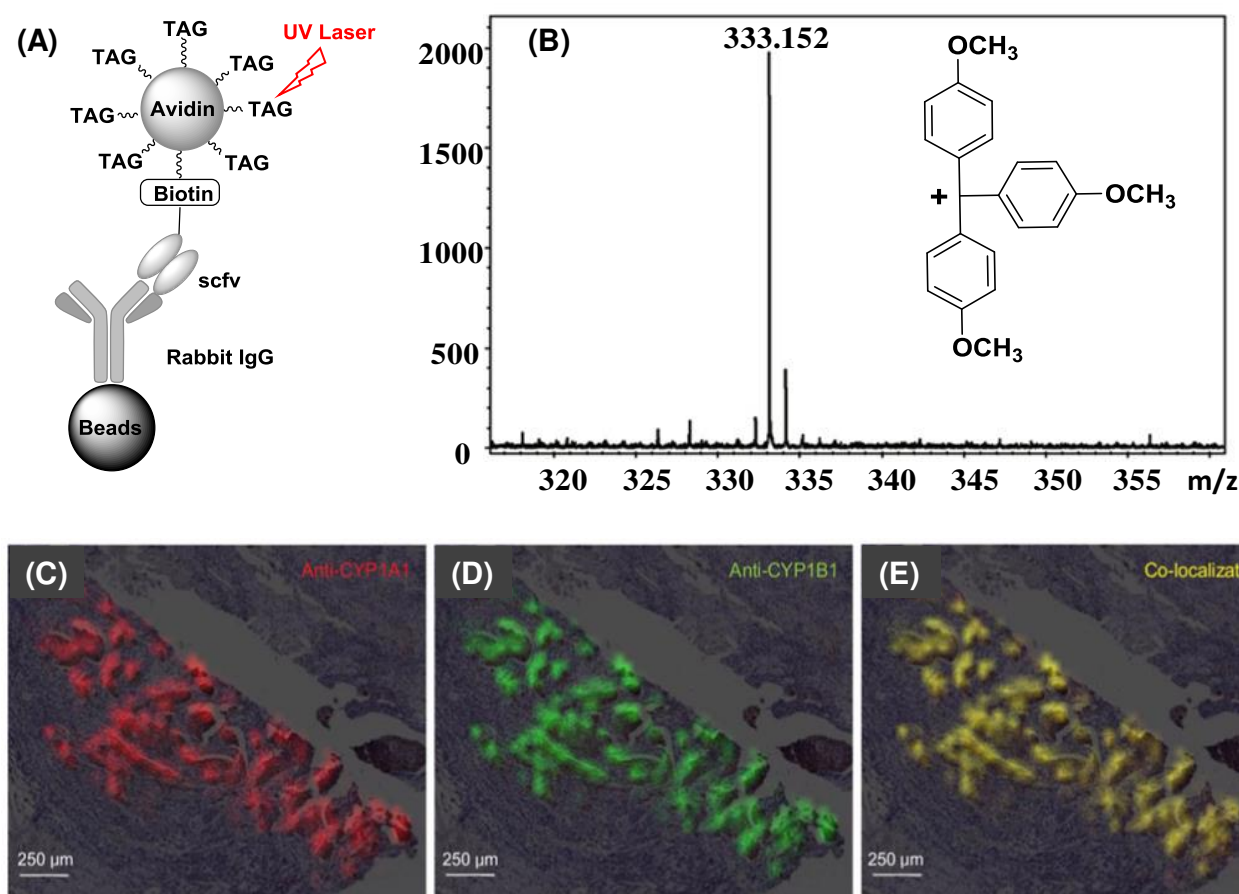
786
 787 **Figure 5.** Concept of TAMSIM to measure protein spatial distributions in tissue sections with MSI using mass tag reporter
 788 ions conjugated via a photocleavable trityl group to antibodies. (A) Schematic representation of the mass-tag reporter ion
 789 released via photodissociation as a result of UV-laser irradiation upon cleavage of the trityl group coupled to the affinity
 790 tag. (B) Reaction steps of the conjugation of a mass tag reporter to an antibody via a photocleavable group and the release
 791 of the mass-tag reporter ion upon UV-laser irradiation. The photocleavable mass-tag reporter reagent contains an NHS
 792 ester as reactive group for covalent attachment to primary amino groups e.g. to the lysine residues of an antibody. In the
 793 ionization interface of the mass spectrometer the trityl groups absorb UV light resulting in the cleavage of the C-S bond
 794 and the release of the ionized mass-tag reporter ion. (C) Improved tags have the structure of alkyl or aromatic groups for
 795 mass tuning and exhibit higher stabilization of residue R on the tag. Plot (A) and (B) were adapted with permission from
 796 Thiery *et al.*¹⁹⁸ and (C) with permission from Thiery *et al.*²⁰¹. Copyright (2007 and 2008) American Chemical Society.

797
 798 Subsequently, this approach was further improved by the same group (**Figure 5C**)²⁰¹. In contrast to
 799 the previous version of TAMSIM, where the mass tags were coupled to secondary antibodies, the
 800 primary antibody is now directly conjugated with the affinity reagent and incubated with the tissue
 801 section. This improvement has the advantage to increase multiplexing as the approach is not limited
 802 by the number of species available for first and secondary antibody pair production. Additionally,
 803 new reporter tags were prepared, which differ from the previous tags at the level of the amide group.
 804 This new class of tags has conjugated alkynes or substituted aromatic groups, which allow for tuning
 805 the mass of the reporter tag and exhibit higher stabilization of the carbocation on the photocleavable
 806 reporter tag. These structural improvements provide more stable reagents, which facilitates handling

807 and sample preparation. The results showed that fewer fragments of the mass tag were observed in
808 the gas phase, which leads to higher sensitivity. The method allowed to analyze FFPE and fresh frozen
809 samples, with the latter having a lower number of artifact peaks in the mass spectra, as these mainly
810 originated from paraffin in FFPE samples. This improved strategy was successfully applied to
811 generate specific mass spectrometric images of three abundant proteins insulin, chromogranin A, and
812 synaptophysin, and the less abundant proteins/peptides calcitonin and somatostatin localized in
813 Langerhans islets²⁰¹.

814 Nevertheless, trityl-based PC-linkers still have several limitations. The highly hydrophobic character
815 of the tagging reagent limits the number of PC-linker/mass tag reporters that can be conjugated to a
816 single antibody, since it reduces the efficiency of the coupling reaction and the aqueous solubility of
817 the resulting conjugates. To overcome this problem, Thiery *et al.* modified TAMSIM by using
818 recombinant single chain variable fragments (scFv) originally designed from monoclonal IgG
819 antibodies labeled with biotin.¹⁹⁵ The biotinylated scFv was coupled to avidin-holding multiple PC-
820 reporter-tags to the biotin moiety to form an immune complex (IC). Essentially, the IC approach
821 allowed the scFv to be linked to mass tags through biotin/avidin coupling and allowed to prepare the
822 IC reagent in two steps, which was subsequently applied to the tissue section. The scFv linked to the
823 reporter tag using this approach was used to specifically and simultaneously detect CYP1A1 and
824 CYP1B1 in breast tumor tissue sections (**Figure 6**). In 2015, Lorey *et al.*²³ presented a new signal
825 detection method for antibody arrays using laser desorption/ionization-mass spectrometry (LDI-MS)
826 based on small, photocleavable reporter molecules. In this work, signal amplification was achieved
827 with a biotin labeled secondary antibody, where biotin is coupled to avidin holding several
828 photocleavable mass-tags. Next, a highly sensitive sandwich assay is performed with immobilized
829 primary antibody capturing prostate specific antigen (PSA) and the secondary antibody labeled with
830 biotin/avidin/reporter-mass tag. This approach allowed to detect PSA in human plasma at clinically
831 relevant concentrations ranging from 2 µg/mL to 200 pg/mL²³. This assay has not been used for MSI
832 yet, but it provides the option to determine the distribution of low abundant proteins in tissue sections.
833 Yang *et al.*²⁰² developed an activity-based MSI approach using reporter mass tags, which provides
834 high spatial resolution, and high sensitivity through the use of signal amplification chemistry and high
835 target specificity (**Figure 7**). In this approach, an activity-based probe (fluorophosphonate) that is
836 specific for serine hydrolases is attached to a dendrimer through click chemistry containing more than
837 900 reporter tags leading to a signal amplification of nearly 3 orders of magnitude. On irradiation of
838 the labeled tissue by the laser beam in a raster pattern, the mass tags are liberated and recorded by the
839 mass spectrometer. Consequently, the ion image of the mass tag reveals the distribution of active
840 serine hydrolases in rat brain and mouse embryo tissue sections. Hong *et al.* reported a mass tag-
841 based MSI method that enables matrix-free MSI of protein biomarkers in FFPE tissues²⁰³. It involves

842 binding of the target protein with a primary antibody, followed by binding with a secondary antibody-
 843 enzyme conjugate. The substrate of the enzyme coupled to the secondary antibody is then added to
 844 the tissue section, and the enzyme converts the substrate to a product, which can be detected by LDI.
 845 The product is deposited at the location of the target protein by precipitation and the precipitates (e.g.
 846 diazonium salts) serve as reporter tags detected by mass spectrometry. The enzymes horseradish
 847 peroxidase and alkaline phosphatase and various substrates have been used to demonstrate the
 848 feasibility of this novel MSI method to image protein targets in FFPE tissue samples. The spatial
 849 resolution of this is only limited by the laser spot size of the commercially available instrument
 850 reaching limit of 10 μm without overlapping laser sampling area.



851

852 **Figure 6.** (A) Structure of the reagent used for targeted detection of reporter-tagged avidin bound to biotinylated A10B
 853 scFv on rabbit IgG coated beads used to optimize scfv-mass tag labeling. (B) Mass spectrum showing the released mass
 854 tag upon UV laser irradiation. (C-E) MS ion image of CYP1A1, CYP1B1 and both compounds on breast cancer tissue
 855 sections obtained by visualizing ion distribution of target compound specific reporter mass tag. The plot (E), which
 856 overlays the red and green colors of CYP1A1 and CYP1B1, respectively, shows that these two compounds are perfectly
 857 co-localized in the same tissue section. Adapted with permission from Thierry *et al.*¹⁹⁵. Copyright (2012) American
 858 Society for Mass Spectrometry.

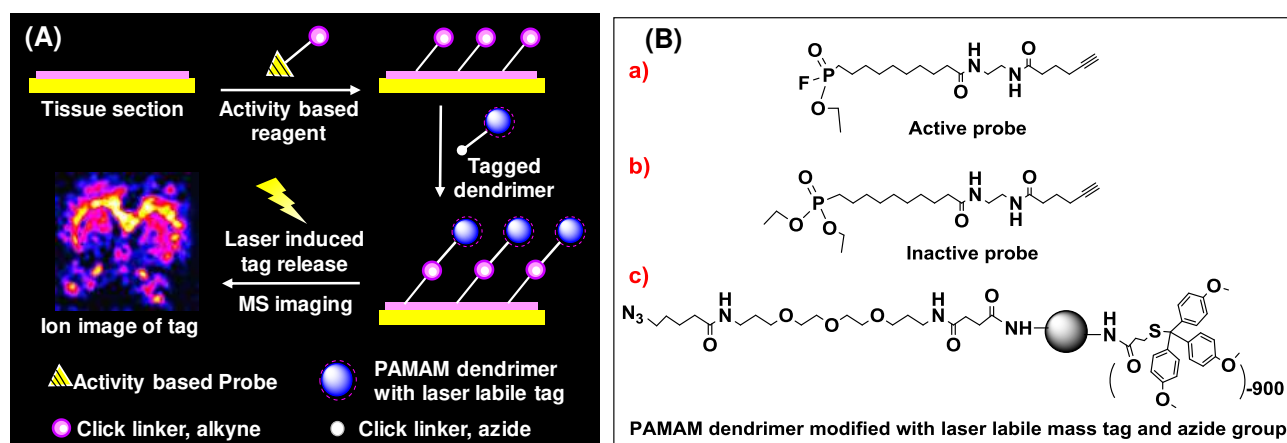


Figure 7. (A) MSI strategy using an activity-based probe conjugated to a PAMAM dendrimer modified with photocleavable mass tags. Structures of (B-a) active probe, (B-b) inactive probe, and (B-c) modified PAMAM dendrimer with the photocleavable mass tag and an azide group used to couple the PAMAM dendrimer with the alkyne group of the activity probe in tissue using click chemistry. Adapted with permission from Yang *et al.*²⁰². Copyright (2012) American Chemical Society.

The PC-linker reporter tag strategy in MSI has significant advantages. The ability to detect a wide variety of proteins without the need of applying matrix helps to overcome previous limitations of MALDI MSI of intact proteins, i.e. low spatial resolution, restriction to the detection of high abundant and low molecular weight proteins with limited dynamic concentration range and incompatibility with FFPE tissues. The mass tag methods can be used to perform MSI on low-abundance proteins or to reveal the localization of active proteins in tissue. This approach can perform highly multiplexed analysis due to its ability to incorporate a large variety of reporter (mass) tags. However, it relies on the quality, cross-reactivity and reliability of the affinity tag (antibody, affimer or affinity probe) and provides a targeted and indirect signal of the proteins of interest, which alleviates to a certain extent the advantage of using mass spectrometry for detection.

4 Conclusions and Perspectives

The methods reviewed here emphasize the immense potential of MSI for studying the spatial distribution of proteins in tissue samples. Major challenges associated with sample preparation, data processing, and MS instrument design have been identified, particularly in order to simultaneously detect the distribution of large numbers of proteins with high spatial resolution and to extend the detected dynamic range with more accurate quantification. MSI of proteins is a rapidly developing field in analytical chemistry and recent developments such as novel ionization techniques, novel strategies for chemical labeling with photocleavable reporter (mass) tags, novel fragmentation approaches, and the improvements in mass spectrometry scanning speed are advancing all aspects of

885 this technology. For example, the mass tag-based LDI MSI approach, implemented as the Tag-Mass
 886 and TAMSIM methods, exhibits significant potential to achieve multiplexed imaging of proteins with
 887 high resolution in tissue sections with important applications in pathology laboratories as it can be
 888 used concurrent with immunohistochemistry staining. Recent advances in top-down mass
 889 spectrometry such as the enhanced transmission of high molecular mass protein ions^{204,205} or the
 890 introduction of novel fragmentation approaches such as UVPD, which allow more complete
 891 fragmentation of intact proteins^{173–175} confidently without the requirement for extensive cleanup, will
 892 further contribute to bringing protein MSI technology to maturation. Another trend holding potential
 893 improvement of protein MSI, is the combination of DESI and MALDI MSI, allowing to measure the
 894 lipid and protein distributions in subsequent analyses in the same tissue section²⁰⁶. In addition, MSI
 895 data can be integrated with spectroscopic images, including automatic annotation transfer of anatomic
 896 structures from microscopic images or from anatomical databases expanding the information content
 897 but also the dimensionality of the data.^{207–209} Combination of anatomical annotation, image fusion
 898 with bioinformatics solutions enabling to process and evaluate the large volume of MSI data in
 899 interactive way without loss of information would further improve the information that can be
 900 obtained from MSI studies.

901 All these technological advances will contribute to the full development of the MSI technology to
 902 profile protein distributions in tissues and will allow to broaden its scope in various fundamental and
 903 clinical applications, including new ways of pathological evaluation of tissue biopsies taken from
 904 patients to support diagnostics.

905 References

- 906 1 S. Rauser, S.-O. Deininger, D. Suckau, H. Höfler and A. Walch, *Expert Rev. Proteomics*, 2010, **7**, 927–
 907 41.
- 908 2 C. Schöne, H. Höfler and A. Walch, *Clin. Biochem.*, 2013, **46**, 539–545.
- 909 3 E. G. Solon, A. Schweitzer, M. Stoeckli and B. Prideaux, *AAPS J.*, 2010, **12**, 11–26.
- 910 4 G. Dimcevski, F. G. Erchinger, R. Havre and O. H. Gilja, *World J. Gastroenterol.*, 2013, **19**, 7247.
- 911 5 P. Flechsig, C. Walker, C. Kratochwil, L. König, A. Iagura, J. Moltz, T. Holland-Letz, H.-U. Kauczor, U.
 912 Haberkorn and F. L. Giesel, *Mol. Imaging Biol.*, 2017, 1–9.
- 913 6 M. F. Dumont, H. A. Hoffman, P. R. S. Yoon, L. S. Conklin, S. R. Saha, J. Paglione, R. W. Sze and R.
 914 Fernandes, *Bioconjug. Chem.*, 2014, **25**, 129–137.
- 915 7 R. Atreya, H. Neumann, C. Neufert, M. J. Waldner, U. Billmeier, Y. Zopf, M. Willma, C. App, T.
 916 Münster, H. Kessler, S. Maas, B. Gebhardt, R. Heimke-Brinck, E. Reuter, F. Dörje, T. T. Rau, W. Uter,
 917 T. D. Wang, R. Kiesslich, M. Vieth, E. Hannappel and M. F. Neurath, *Nat. Med.*, 2014, **20**, 313–318.
- 918 8 K. Cheng, S. R. Kothapalli, H. Liu, A. L. Koh, J. V. Jokerst, H. Jiang, M. Yang, J. Li, J. Levi, J. C. Wu, S. S.
 919 Gambhir and Z. Cheng, *J. Am. Chem. Soc.*, 2014, **136**, 3560–3571.
- 920 9 G. Lu and B. Fei, *J. Biomed. Opt.*, 2014, **19**, 010901.

- 921 10 W. R. W. Martin, *Mol. Imaging Biol.*, 2007, **9**, 196–203.
- 922 11 O. Golf, N. Strittmatter, T. Karancsi, S. D. Pringle, A. V. M. Speller, A. Mroz, J. M. Kinross, N. Abbassi-
923 Ghadi, E. A. Jones and Z. Takats, *Anal. Chem.*, 2015, **87**, 2527–2534.
- 924 12 J. H. Gross, *Anal. Bioanal. Chem.*, 2014, **406**, 63–80.
- 925 13 R. M. Alberici, P. H. Vendramini and M. N. Eberlin, *Anal. Methods*, 2017, **9**, 5029–5036.
- 926 14 E. Esquenazi, C. Coates, L. Simmons, D. Gonzalez, W. H. Gerwick and P. C. Dorrestein, *Mol. Biosyst.*,
927 2008, **4**, 562–570.
- 928 15 Y. J. Lee, D. C. Perdian, Z. Song, E. S. Yeung and B. J. Nikolau, *Plant J.*, 2012, **70**, 81–95.
- 929 16 A. Römpf, S. Guenther, Y. Schober, O. Schulz, Z. Takats, W. Kummer and B. Spengler, *Angew.*
930 *Chemie - Int. Ed.*, 2010, **49**, 3834–3838.
- 931 17 K. Huber, A. Feuchtinger, D. M. Borgmann, Z. Li, M. Aichler, S. M. Hauck, H. Zitzelsberger, M.
932 Schwaiger, U. Keller and A. Walch, *Anal. Chem.*, 2014, **86**, 10568–10575.
- 933 18 R. Vismeh, D. J. Waldon, Y. Teffera and Z. Zhao, *Anal. Chem.*, 2012, **84**, 5439–5445.
- 934 19 M. R. Groseclose and S. Castellino, *Anal. Chem.*, 2013, **85**, 10099–10106.
- 935 20 J. G. Swales, J. W. Tucker, N. Strittmatter, A. Nilsson, D. Cobice, M. R. Clench, C. L. Mackay, P. E.
936 Andren, Z. Takáts, P. J. H. Webborn and R. J. A. Goodwin, *Anal. Chem.*, 2014, **86**, 8473–80.
- 937 21 S. J. B. Dunham, J. F. Ellis, B. Li and J. V. Sweedler, *Acc. Chem. Res.*, 2017, **50**, 96–104.
- 938 22 R. Lemaire, J. Stauber, M. Wisztorski, C. Van Camp, A. Desmons, M. Deschamps, G. Proess, I. Rudlof,
939 A. S. Woods, R. Day, M. Salzert and I. Fournier, *J. Proteome Res.*, 2007, **13**, 2057–2067.
- 940 23 M. Lorey, B. Adler, H. Yan, R. Soliymani, S. Ekström, J. Yli-Kauhaluoma, T. Laurell and M. Baumann,
941 *Anal. Chem.*, 2015, **87**, 5255–5262.
- 942 24 J. Stauber, M. El Ayed, M. Wisztorski, M. Salzert and I. Fournier, Humana Press, Totowa, NJ, 2010, pp.
943 339–361.
- 944 25 L. M. Cole, K. Mahmoud, S. Haywood-Small, G. M. Tozer, D. P. Smith and M. R. Clench, *Rapid*
945 *Commun. Mass Spectrom.*, 2013, **27**, 2355–2362.
- 946 26 K. Dreisewerd, *Anal. Bioanal. Chem.*, 2014, **406**, 2261–2278.
- 947 27 B. Balluff, C. Schöne, H. Höfler and A. Walch, *Histochem. Cell Biol.*, 2011, **136**, 227–44.
- 948 28 B. K. Kaletas, I. M. Van Der Wiel, J. Stauber, L. J. Dekker, C. Güzel, J. M. Kros, T. M. Luider and R. M.
949 A. Heeren, *Proteomics*, 2009, **9**, 2622–2633.
- 950 29 R. J. A. Goodwin, *J. Proteomics*, 2012, **75**, 4893–4911.
- 951 30 M.-C. Djidja, E. Claude, P. Scriven, D. W. Allen, V. A. Carolan and M. R. Clench, *Biochim. Biophys. Acta*
952 *- Proteins Proteomics*, 2017, **1865**, 901–906.
- 953 31 S. Ongay, M. Langelaar-Makkinje, M. P. Stoop, N. Liu, H. Overkleeft, T. M. Luider, G. M. M. Groothuis
954 and R. Bischoff, *ChemBioChem*, 2018, **19**, 736–743.
- 955 32 R. J. A. Goodwin, S. R. Pennington and A. R. Pitt, *Proteomics*, 2008, **8**, 3785–3800.
- 956 33 P. Chaurand, J. L. Norris, D. S. Cornett, J. A. Mobley and R. M. Caprioli, *J. Proteome Res.*, 2006, **5**,
957 2889–2900.
- 958 34 R. J. A. Goodwin, S. R. Pennington and A. R. Pitt, *Proteomics*, 2008, **8**, 3785–3800.
- 959 35 W. Bouschen and B. Spengler, *Int. J. Mass Spectrom.*, 2007, **266**, 129–137.
- 960 36 T. W. Jaskolla, M. Karas, U. Roth, K. Steinert, C. Menzel and K. Reihs, *J. Am. Soc. Mass Spectrom.*,
961 2009, **20**, 1104–1114.
- 962 37 P. Chaurand, S. A. Schwartz, D. Billheimer, B. J. Xu, A. Crecelius and R. M. Caprioli, *Anal. Chem.*, 2004,
963 **76**, 1145–1155.

964 38 H. R. Aerni, D. S. Cornett and R. M. Caprioli, *Anal. Chem.*, 2006, **78**, 827–834.

965 39 H. Liebl, *J. Appl. Phys.*, 1967, **38**, 5277–5283.

966 40 H. Nygren and P. Malmberg, *Proteomics*, 2010, **10**, 1694–1698.

967 41 E. de. Hoffmann and V. Stroobant, *Mass spectrometry : principles and applications*, J. Wiley, 2007.

968 42 J. J. D. Fitzgerald, P. Kunnath and A. V Walker, *Technology*, 2010, **82**, 4413–4419.

969 43 J. Xu, S. Ostrowski, C. Szakal, A. G. Ewing and N. Winograd, in *Applied Surface Science*, 2004, vol.
970 231–232, pp. 159–163.

971 44 R. M. A. Heeren, L. A. McDonnell, E. Amstalden, S. L. Luxembourg, A. F. M. Altelaar and S. R. Piersma,
972 *Appl. Surf. Sci.*, 2006, **252**, 6827–6835.

973 45 S. G. Boxer, M. L. Kraft and P. K. Weber, *Annu. Rev. Biophys.*, 2009, **38**, 53–74.

974 46 F. Kollmer, W. Paul, M. Krehl and E. Niehuis, in *Surface and Interface Analysis*, 2013, vol. 45, pp.
975 312–314.

976 47 N. Winograd, *Anal. Chem.*, 2015, **87**, 328–333.

977 48 J. S. Fletcher and J. C. Vickerman, *Anal. Bioanal. Chem.*, 2010, **396**, 85–104.

978 49 J. S. Fletcher, J. C. Vickerman and N. Winograd, *Curr. Opin. Chem. Biol.*, 2011, **15**, 733–740.

979 50 R. Hill, P. Blenkinsopp, S. Thompson, J. Vickerman and J. S. Fletcher, in *Surface and Interface*
980 *Analysis*, 2011, vol. 43, pp. 506–509.

981 51 J. S. Fletcher, N. P. Lockyer, S. Vaidyanathan and J. C. Vickerman, *Anal. Chem.*, 2007, **79**, 2199–2206.

982 52 J. S. Fletcher, N. P. Lockyer and J. C. Vickerman, *Mass Spectrom. Rev.*, 2011, **30**, 142–174.

983 53 E. J. Lanni, S. S. Rubakhin and J. V. Sweedler, *J. Proteomics*, 2012, **75**, 5036–5051.

984 54 S. Solé-Domènech, P. Sjövall, V. Vukojević, R. Fernando, A. Codita, S. Salve, N. Bogdanović, A. H.
985 Mohammed, P. Hammarström, K. P. R. Nilsson, F. M. Laferla, S. Jacob, P. O. Berggren, L. Giménez-
986 Llort, M. Schalling, L. Terenius and B. Johansson, *Acta Neuropathol.*, 2013, **125**, 145–157.

987 55 J. S. Fletcher, *Analyst*, 2009, **134**, 2204.

988 56 J. S. Fletcher, S. Rabbani, A. Henderson, P. Blenkinsopp, S. P. Thompson, N. P. Lockyer and J. C.
989 Vickerman, *Anal. Chem.*, 2008, **80**, 9058–9064.

990 57 E. J. Lanni, S. J. B. Dunham, P. Nemes, S. S. Rubakhin and J. V. Sweedler, *J. Am. Soc. Mass Spectrom.*,
991 2014, **25**, 1897–1907.

992 58 K. Chughtai and R. M. A. Heeren, *Chem. Rev.*, 2010, **110**, 3237–3277.

993 59 Z. Takáts, J. M. Wiseman, B. Gologan and R. G. Cooks, *Science (80-.)*, 2004, **306**, 471–473.

994 60 J. M. Wiseman, D. R. Ifa, Q. Song and R. G. Cooks, *Angew. Chemie - Int. Ed.*, 2006, **45**, 7188–7192.

995 61 M. Nefliu, J. N. Smith, A. Venter and R. G. Cooks, *J. Am. Soc. Mass Spectrom.*, 2008, **19**, 420–427.

996 62 C. Wu, D. R. Ifa, N. E. Manicke and R. G. Cooks, *Analyst*, 2010, **135**, 28–32.

997 63 T. Müller, S. Oradu, D. R. Ifa, R. G. Cooks and B. Kräutler, *Anal. Chem.*, 2011, **83**, 5754–5761.

998 64 L. S. Eberlin, C. R. Ferreira, A. L. Dill, D. R. Ifa and R. G. Cooks, *Biochim. Biophys. Acta - Mol. Cell Biol.*
999 *Lipids*, 2011, **1811**, 946–960.

1000 65 S. R. Ellis, C. Wu, J. M. Deeley, X. Zhu, R. J. W. Truscott, M. in het Panhuis, R. G. Cooks, T. W. Mitchell
1001 and S. J. Blanksby, *J. Am. Soc. Mass Spectrom.*, 2010, **21**, 2095–2104.

1002 66 J. Thunig, S. H. Hansen and C. Janfelt, *Anal. Chem. (Washington, DC, United States)*, 2011, **83**, 3256–
1003 3259.

1004 67 J. Watrous, N. Hendricks, M. Meehan and P. C. Dorrestein, *Anal. Chem.*, 2010, **82**, 1598–1600.

1005 68 Z. Takáts, J. M. Wiseman and R. G. Cooks, *J. Mass Spectrom.*, 2005, **40**, 1261–1275.

1006 69 C. Janfelt and A. W. Nørgaard, *J. Am. Soc. Mass Spectrom.*, 2012, **23**, 1670–1678.

1007 70 M. Girod, Y. Shi, J. X. Cheng and R. G. Cooks, *Anal. Chem.*, 2011, **83**, 207–215.

1008 71 V. Kertesz and G. J. Van Berkel, *Rapid Commun. Mass Spectrom.*, 2008, **22**, 2639–2644.

1009 72 W. Rao, D. J. Scurr, J. Burstson, M. R. Alexander and D. a Barrett, *Analyst*, 2012, **137**, 3946.

1010 73 K. Y. Garza, C. L. Feider, D. R. Klein, J. A. Rosenberg, J. S. Brodbelt and L. S. Eberlin, *Anal. Chem.*,
1011 2018, **90**, 7785–7789.

1012 74 P. Nemes and A. Vertes, *Anal. Chem.*, 2007, **79**, 8098–8106.

1013 75 J. Zou, F. Talbot, A. Tata, L. Ermini, K. Franjic, M. Ventura, J. Zheng, H. Ginsberg, M. Post, D. R. Ifa, D.
1014 Jaffray, R. J. D. Miller and A. Zarrine-Afsar, *Anal. Chem.*, 2015, **87**, 12071–12079.

1015 76 M. Stoeckli, P. Chaurand, D. E. Hallahan and R. M. Caprioli, *Nat. Med.*, 2001, **7**, 493–496.

1016 77 M. Karas and F. Hillenkamp, *Anal. Chem.*, 1988, **60**, 2299–2301.

1017 78 R. Limberg and S. D. Gmbh, 1988, **2301**, 2299–2301.

1018 79 A. Römpf and B. Spengler, *Histochem. Cell Biol.*, 2013, **139**, 759–783.

1019 80 R. M. Caprioli, T. B. Farmer and J. Gile, *Anal. Chem.*, 1997, **69**, 4751–60.

1020 81 P. Chaurand, S. A. Schwartz and R. M. Caprioli, *Curr. Opin. Chem. Biol.*, 2002, **6**, 676–681.

1021 82 N. Verbeeck, J. Yang, B. De Moor, R. M. Caprioli, E. Waelkens and R. Van de Plas, *Anal. Chem.*, 2014,
1022 **86**, 8974–82.

1023 83 M. Karas and R. Kr??ger, *Chem. Rev.*, 2003, **103**, 427–439.

1024 84 J. Yang and R. M. Caprioli, *Anal. Chem.*, 2011, **83**, 5728–5734.

1025 85 S. R. Ellis, J. H. Jungmann, D. F. Smith, J. Soltwisch and R. M. a Heeren, *Angew. Chem. Int. Ed. Engl.*,
1026 2013, **52**, 11261–4.

1027 86 J. E. Yarnold, B. R. Hamilton, D. T. Welsh, G. F. Pool, D. J. Venter and A. R. Carroll, *Mol. Biosyst.*, 2012,
1028 **8**, 2249–59.

1029 87 P. M. Angel, J. M. Spraggins, H. S. Baldwin and R. Caprioli, *Anal. Chem.*, 2012, **84**, 1557–1564.

1030 88 M. Poetzsch, A. E. Steuer, A. T. Roemmelt, M. R. Baumgartner and T. Kraemer, *Anal. Chem.*, 2014,
1031 **86**, 11758–65.

1032 89 X. C. Yu, O. V. Borisov, M. Alvarez, D. A. Michels, Y. J. Wang and V. Ling, *Anal. Chem.*, 2009, **81**, 9282–
1033 9290.

1034 90 A. Römpf, S. Guenther, Z. Takats and B. Spengler, *Anal. Bioanal. Chem.*, 2011, **401**, 65–73.

1035 91 C. Dai, L. H. Cazares, L. Wang, Y. Chu, S. L. Wang, D. a Troyer, O. J. Semmes, R. R. Drake and B. Wang,
1036 *Chem. Commun. (Camb).*, 2011, **47**, 10338–10340.

1037 92 P. Chaurand, K. E. Schriver and R. M. Caprioli, *J. Mass Spectrom.*, 2007, **42**, 476–489.

1038 93 Y. Li, B. Shrestha and A. Vertes, *Anal. Chem.*, 2008, **80**, 407–420.

1039 94 Y. Li, B. Shrestha and A. Vertes, 2007, **79**, 523–532.

1040 95 D. R. am Bhandari, M. Schott, A. Römpf, A. Vilcinskas and B. Spengler, *Anal. Bioanal. Chem.*, 2015,
1041 **407**, 2189–2201.

1042 96 S. M. Khalil, A. Römpf, J. Pretzel, K. Becker and B. Spengler, *Anal. Chem.*, 2015, **87**, 11309–11316.

1043 97 M. Kompauer, S. Heiles and B. Spengler, *Nat. Methods*, 2016, **14**, 90–96.

1044 98 C. Keller, J. Maeda, D. Jayaraman, S. Chakraborty, M. R. Sussman, J. M. Harris, J.-M. Ané and L. Li,
1045 *Front. Plant Sci.*, 2018, **9**, 1238.

1046 99 J. S. Sampson, A. M. Hawkridge and D. C. Muddiman, *J. Am. Soc. Mass Spectrom.*, 2006, **17**, 1712–
1047 1716.

1048 100 J. S. Sampson and D. C. Muddiman, *Rapid Commun. Mass Spectrom.*, 2009, **23**, 1989–1992.

1049 101 J. Sabine Becker, *J. Mass Spectrom.*, 2013, **48**, 255–268.

1050 102 J. S. Becker, M. Zoriy, A. Matusch, B. Wu, D. Salber, C. Palm and J. S. Becker, *Mass Spectrom. Rev.*,
1051 2010, **29**, 156–175.

1052 103 I. Feldmann, C. U. Koehler, P. H. Roos and N. Jakubowski, *J. Anal. At. Spectrom.*, 2006, **21**, 1006.

1053 104 F. Schueder, J. Lara-Gutiérrez, B. J. Beliveau, S. K. Saka, H. M. Sasaki, J. B. Woehrstein, M. T. Strauss,
1054 H. Grabmayr, P. Yin and R. Jungmann, *Nat. Commun.*, 2017, **8**, 2090.

1055 105 J. R. Chirinos, D. D. Oropeza, J. J. Gonzalez, H. Hou, M. Morey, V. Zorba and R. E. Russo, *J. Anal. At.*
1056 *Spectrom.*, 2014, **29**, 1292–1298.

1057 106 S. K. EDELMAN, *Echocardiography*, 1988, **5**, 269–272.

1058 107 B. Spengler and M. Hubert, *J. Am. Soc. Mass Spectrom.*, 2002, **13**, 735–748.

1059 108 D. Trede, S. Schiffler, M. Becker, S. Wirtz, K. Steinhorst, J. Strehlow, M. Aichler, J. H. Kobarg, J.
1060 Oetjen, A. Dyatlov, S. Heldmann, A. Walch, H. Thiele, P. Maass and T. Alexandrov, *Anal. Chem.*, 2012,
1061 **84**, 6079–6087.

1062 109 M. Andersson, M. R. Groseclose, A. Y. Deutch and R. M. Caprioli, *Nat. Methods*, 2008, **5**, 101–108.

1063 110 L. Morosi, S. Giordano, F. Falcetta, R. Frapolli, S. A. Licandro, C. Matteo, M. Zucchetti, P. Ubezio, E.
1064 Erba, S. Visentin, M. D’Incalci and E. Davoli, *Clin. Pharmacol. Ther.*, 2017, **102**, 748–751.

1065 111 E. H. Seeley and R. M. Caprioli, *Anal. Chem.*, 2012, **84**, 2105–2110.

1066 112 S. Giordano, L. Morosi, P. Veglianese, S. A. Licandro, R. Frapolli, M. Zucchetti, G. Cappelletti, L.
1067 Falcicola, V. Pifferi, S. Visentin, M. D’Incalci and E. Davoli, *Sci. Rep.*, 2016, **6**, 37027.

1068 113 D. Touboul and A. Brunelle, *Mass Spectrometry Imaging of Small Molecules*, Springer New York, New
1069 York, NY, 2015, vol. 1203.

1070 114 G. Jacobs, *Prof. Surv. Mag.*

1071 115 S. Keren, C. Zavaleta, Z. Cheng, A. de la Zerda, O. Gheysens and S. S. Gambhir, *Proc. Natl. Acad. Sci.*
1072 *U. S. A.*, 2008, **105**, 5844–9.

1073 116 J. C. Jurchen, S. S. Rubakhin and J. V. Sweedler, *J. Am. Soc. Mass Spectrom.*, 2005, **16**, 1654–1659.

1074 117 M. Nazari and D. C. Muddiman, *Anal. Bioanal. Chem.*, 2015, **407**, 2265–2271.

1075 118 A. D. Palmer and T. Alexandrov, *Anal. Chem.*, 2015, **87**, 4055–4062.

1076 119 D. Trede, S. Schi, M. Becker, S. Wirtz, K. Steinhorst, J. Strehlow, M. Aichler, J. H. Kobarg, J. Oetjen, A.
1077 Dyatlov, S. Heldmann, A. Walch, H. Thiele, P. Maass and T. Alexandrov, *Anal. Chem.*, 2012, **84**, 6079–
1078 6087.

1079 120 A. Bouslimani, C. Porto, C. M. Rath, M. Wang, Y. Guo, A. Gonzalez, D. Berg-Lyon, G. Ackermann, G. J.
1080 Moeller Christensen, T. Nakatsuji, L. Zhang, A. W. Borkowski, M. J. Meehan, K. Dorrestein, R. L. Gallo,
1081 N. Bandeira, R. Knight, T. Alexandrov and P. C. Dorrestein, *Proc. Natl. Acad. Sci.*, 2015, **112**, E2120–
1082 E2129.

1083 121 T. Alexandrov, *BMC Bioinformatics*, 2012, **13 Suppl 1**, S11.

1084 122 E. E. Jones, C. Quiason, S. Dale and S. K. Shahidi-Latham, *J. Am. Soc. Mass Spectrom.*, 2017, **28**,
1085 1709–1715.

1086 123 F. Suits, T. E. Fehniger, Á. Végvári, G. Marko-Varga and P. Horvatovich, *Anal. Chem.*, 2013, **85**, 4398–
1087 4404.

1088 124 T. E. Fehniger, F. Suits, Á. Végvári, P. Horvatovich, M. Foster and G. Marko-Varga, *Proteomics*, 2014,
1089 **14**, 862–871.

1090 125 W. M. Abdelmoula, K. Škrášková, B. Balluff, R. J. Carreira, E. A. Tolner, B. P. F. Lelieveldt, L. Van Der
1091 Maaten, H. Morreau, A. M. J. M. Van Den Maagdenberg, R. M. A. Heeren, L. A. McDonnell and J.

1092 Dijkstra, *Anal. Chem.*, 2014, **86**, 9204–9211.

1093 126 L. A. McDonnell, A. Walch, M. Stoeckli and G. L. Corthals, *J. Proteome Res.*, 2014, **13**, 1138–1142.

1094 127 T. Alexandrov, *BMC Bioinformatics*, 2012, **13 Suppl 1**, S11.

1095 128 T. Schramm, A. Hester, I. Klinkert, J. P. Both, R. M. A. Heeren, A. Brunelle, O. Laprévotte, N.
1096 Desbenoit, M. F. Robbe, M. Stoeckli, B. Spengler and A. Römpf, *J. Proteomics*, 2012, **75**, 5106–5110.

1097 129 A. Römpf, T. Schramm, A. Hester, I. Klinkert, J.-P. Both, R. M. A. Heeren, M. Stöckli and B. Spengler,
1098 *Methods Mol. Biol.*, 2011, **696**, 205–24.

1099 130 G. Mayer, L. Montecchi-Palazzi, D. Ovelheiro, A. R. Jones, P. A. Binz, E. W. Deutsch, M. Chambers, M.
1100 Kallhardt, F. Levander, J. Shofstahl, S. Orchard, J. A. Vizcaíno, H. Hermjakob, C. Stephan, H. E. Meyer
1101 and M. Eisenacher, *Database*, 2013, **2013**, bat009-bat009.

1102 131 O. Jardin-Mathé, D. Bonnel, J. Franck, M. Wisztorski, E. Macagno, I. Fournier and M. Salzter, *J.*
1103 *Proteomics*, 2008, **71**, 332–345.

1104 132 M. F. Robbe, J. P. Both, B. Prideaux, I. Klinkert, V. Picaud, T. Schramm, A. Hester, V. Guevara, M.
1105 Stoeckli, A. Roempp, R. Heeren, B. Spengler, O. Gala and S. Haana, *Eur. J. Mass Spectrom.*, 2014, **20**,
1106 351–360.

1107 133 J. Oetjen, K. Veselkov, J. Watrous, J. S. McKenzie, M. Becker, L. Hauberg-Lotte, J. H. Kobarg, N.
1108 Strittmatter, A. K. Mróz, F. Hoffmann, D. Trede, A. Palmer, S. Schiffler, K. Steinhorst, M. Aichler, R.
1109 Goldin, O. Guntinas-Lichius, F. von Eggeling, H. Thiele, K. Maedler, A. Walch, P. Maass, P. C.
1110 Dorrestein, Z. Takats and T. Alexandrov, *Gigascience*, 2015, **4**, 20.

1111 134 V. Delcourt, J. Franck, E. Leblanc, F. Narducci, Y.-M. Robin, J.-P. Gimeno, J. Quanco, M. Wisztorski, F.
1112 Kobeissy, J.-F. Jacques, X. Roucou, M. Salzter and I. Fournier, *EBioMedicine*, 2017, **21**, 55–64.

1113 135 J. Sarsby, N. J. Martin, P. F. Lalor, J. Bunch and H. J. Cooper, *J. Am. Soc. Mass Spectrom.*, 2014, **25**,
1114 1953–1961.

1115 136 M. Lagarrigue, M. Becker, R. Lavigne, S.-O. Deininger, A. Walch, F. Aubry, D. Suckau and C. Pineau,
1116 *Mol. Cell. Proteomics*, 2011, **10**, M110.005991-M110.005991.

1117 137 M. Ronci, S. Sharma, S. Martin, J. E. Craig and N. H. Voelcker, *J. Proteomics*, 2013, **82**, 27–34.

1118 138 B. D. Leinweber, G. Tsaprailis, T. J. Monks and S. S. Lau, *J. Am. Soc. Mass Spectrom.*, 2009, **20**, 89–95.

1119 139 A. C. Grey, P. Chaurand, R. M. Caprioli and K. L. Schey, *J. Proteome Res.*, 2009, **8**, 3278–3283.

1120 140 J. Franck, R. Longuespée, M. Wisztorski, A. Van Remoortere, R. Van Zeijl, A. Deelder, M. Salzter, L.
1121 McDonnell and I. Fournier, *Med. Sci. Monit.*, 2010, **16**, BR293-9.

1122 141 V. Mainini, G. Bovo, C. Chinello, E. Gianazza, M. Grasso, G. Cattoretti and F. Magni, *Mol. Biosyst.*,
1123 2013, **9**, 1101.

1124 142 D. C. Reiber, T. a Grover and R. S. Brown, *Anal. Chem.*, 1998, **70**, 673–683.

1125 143 N. L. Kelleher, *Anal. Chem.*, 2004, **76**, 196 A-203 A.

1126 144 R. Lemaire, A. Desmons, J. C. Tabet, R. Day, M. Salzter and I. Fournier, *J. Proteome Res.*, 2007, **6**,
1127 1295–1305.

1128 145 R. J. Rose, E. Damoc, E. Denisov, A. Makarov and A. J. R. Heck, *Nat. Methods*, 2012, **9**, 1084–1086.

1129 146 A. van Remoortere, R. J. M. van Zeijl, N. van den Oever, J. Franck, R. Longuespée, M. Wisztorski, M.
1130 Salzter, A. M. Deelder, I. Fournier and L. A. McDonnell, *J. Am. Soc. Mass Spectrom.*, 2010, **21**, 1922–
1131 1929.

1132 147 J. H. Jungmann, L. MacAleese, J. Visser, M. J. J. Vrakking and R. M. A. Heeren, *Anal. Chem.*, 2011, **83**,
1133 7888–7894.

1134 148 R. Knochenmuss and R. Zenobi, *Chem. Rev.*, 2003, **103**, 441–452.

1135 149 K. Dreisewerd, M. Schürenberg, M. Karas and F. Hillenkamp, *Int. J. Mass Spectrom. Ion Process.*,
1136 1995, **141**, 127–148.

1137 150 R. Knochenmuss and L. V. Zhigilei, *Anal. Bioanal. Chem.*, 2012, **402**, 2511–2519.

1138 151 R. Zenobi and R. Knochenmuss, 1999, 337–366.

1139 152 M. Martin-Lorenzo, B. Balluff, A. Sanz-Maroto, R. J. M. van Zeijl, F. Vivanco, G. Alvarez-Llamas and L.
1140 A. McDonnell, *J. Proteomics*, 2014, **108**, 465–468.

1141 153 F. Deutschens, J. Yang and R. M. Caprioli, *J. Mass Spectrom.*, 2011, **46**, 568–571.

1142 154 Y. Schober, S. Guenther, B. Spengler and A. Römpf, *Anal. Chem.*, 2012, **84**, 6293–6297.

1143 155 A. Zavalin, E. M. Todd, P. D. Rawhouser, J. Yang, J. L. Norris and R. M. Caprioli, *J. Mass Spectrom.*,
1144 2012, **47**, 1473–1481.

1145 156 A. Zavalin, J. Yang, K. Hayden, M. Vestal and R. M. Caprioli, *Anal. Bioanal. Chem.*, 2015, **407**, 2337–
1146 42.

1147 157 A. Zavalin, J. Yang and R. Caprioli, *J. Am. Soc. Mass Spectrom.*, 2013, **24**, 1153–1156.

1148 158 M. Koestler, D. Kirsch, A. Hester, A. Leisner, S. Guenther and B. Spengler, *Rapid Commun. Mass
1149 Spectrom.*, 2008, **22**, 3275–3285.

1150 159 R. Van De Plas, J. Yang, J. Spraggins and R. M. Caprioli, *Nat. Methods*, 2015, **12**, 366–372.

1151 160 J. M. Spraggins, D. G. Rizzo, J. L. Moore, M. J. Noto, E. P. Skaar and R. M. Caprioli, *Proteomics*, 2016,
1152 **16**, 1678–1689.

1153 161 D. Calligaris, C. Villard and D. Lafitte, *J. Proteomics*, 2011, **74**, 920–934.

1154 162 D. Bonnel, R. Longuespée, J. Franck, M. Roudbaraki, P. Gosset, R. Day, M. Salzert and I. Fournier,
1155 *Anal. Bioanal. Chem.*, 2011, **401**, 149–165.

1156 163 D. Debois, V. Bertrand, L. Quinton, M. C. De Pauw-Gillet and E. De Pauw, *Anal. Chem.*, 2010, **82**,
1157 4036–4045.

1158 164 J. Hardouin, *Mass Spectrom. Rev.*, 2007, **26**, 672–82.

1159 165 K. E. Burnum, S. L. Frappier and R. M. Caprioli, *Annu. Rev. Anal. Chem. (Palo Alto. Calif.)*, 2008, **1**,
1160 689–705.

1161 166 M. Takayama, *J. Mass Spectrom. Soc. Jpn.*, 2002, **50**, 304–310.

1162 167 P. Chaurand, B. B. DaGue, S. Ma, S. Kasper and R. M. Caprioli, 2001, **40**, 9725–9733.

1163 168 C. S. Raska, C. E. Parker, C. Huang, J. Han, G. L. Glish, M. Pope and C. H. Borchers, *J. Am. Soc. Mass
1164 Spectrom.*, 2002, **13**, 1034–1041.

1165 169 D. Suckau, A. Resemann, M. Schuerenberg, P. Hufnagel, J. Franzen and A. Holle, *Anal. Bioanal.
1166 Chem.*, 2003, **376**, 952–965.

1167 170 T. A. Zimmerman, D. Debois, G. Mazzucchelli, V. Bertrand, M. C. De Pauw-Gillet and E. De Pauw,
1168 *Anal. Chem.*, 2011, **83**, 6090–6097.

1169 171 D. Calligaris, R. Longuespée, D. Debois, D. Asakawa, A. Turtoi, V. Castronovo, A. Noël, V. Bertrand, M.
1170 C. De Pauw-Gillet and E. De Pauw, *Anal. Chem.*, 2013, **85**, 2117–2126.

1171 172 S. Tamara, A. Dyachenko, K. L. Fort, A. A. Makarov, R. A. Scheltema and A. J. R. Heck, *J. Am. Chem.
1172 Soc.*, 2016, **138**, 10860–10868.

1173 173 T. P. Cleland, C. J. DeHart, R. T. Fellers, A. J. VanNispen, J. B. Greer, R. D. LeDuc, W. R. Parker, P. M.
1174 Thomas, N. L. Kelleher and J. S. Brodbelt, *J. Proteome Res.*, 2017, **16**, 2072–2079.

1175 174 S. M. Greer and J. S. Brodbelt, *J. Proteome Res.*, 2018, **17**, 1138–1145.

1176 175 X. Dang and N. L. Young, *Proteomics*, 2014, **14**, 1128–1129.

1177 176 P. Mallick, M. Schirle, S. S. Chen, M. R. Flory, H. Lee, D. Martin, J. Ranish, B. Raught, R. Schmitt, T.
1178 Werner, B. Kuster and R. Aebersold, *Nat. Biotechnol.*, 2007, **25**, 125–31.

1179 177 M. R. Groseclose, M. Andersson, W. M. Hardesty and R. M. Caprioli, *J. Mass Spectrom.*, 2007, **42**,
1180 254–262.

1181 178 M. Ronci, S. Sharma, T. Chataway, K. P. Burdon, S. Martin, J. E. Craig and N. H. Voelcker, *J. Proteome*
1182 *Res.*, 2011, **10**, 3522–3529.

1183 179 H. C. Diehl, B. Beine, J. Elm, D. Trede, M. Ahrens, M. Eisenacher, K. Marcus, H. E. Meyer and C.
1184 Henkel, *Anal. Bioanal. Chem.*, 2015, **407**, 2223–2243.

1185 180 D. E. Clemmer and M. F. Jarrold, *J. Mass Spectrom.*, 1997, **32**, 577–592.

1186 181 M. C. Djidja, E. Claude, M. F. Snel, P. Scriven, S. Francese, V. Carolan and M. R. Clench, *J. Proteome*
1187 *Res.*, 2009, **8**, 4876–4884.

1188 182 J. Stauber, L. MacAleese, J. Franck, E. Claude, M. Snel, B. K. Kaletas, I. M. V. D. Wiel, M. Wisztorski, I.
1189 Fournier and R. M. a Heeren, *J. Am. Soc. Mass Spectrom.*, 2010, **21**, 338–47.

1190 183 Y. Schober, T. Schramm, B. Spengler and A. Römpf, *Rapid Commun. Mass Spectrom.*, 2011, **25**,
1191 2475–2483.

1192 184 B. Heijs, R. J. Carreira, E. A. Tolner, A. H. De Ru, A. M. J. M. Van Den Maagdenberg, P. A. Van Veelen
1193 and L. A. McDonnell, *Anal. Chem.*, 2015, **87**, 1867–1875.

1194 185 J. Franck, M. El Ayed, M. Wisztorski, M. Salzter and I. Fournier, *Anal. Chem.*, 2009, **81**, 8305–8317.

1195 186 P. J. Brownridge, V. M. Harman, D. M. Simpson and R. J. Beynon, in *Methods in molecular biology*
1196 *(Clifton, N.J.)*, 2012, vol. 893, pp. 267–293.

1197 187 T. Porta, A. Lesur, E. Varesio and G. Hopfgartner, *Anal. Bioanal. Chem.*, 2015, **407**, 2177–2187.

1198 188 M. Komatsu, Y. Murayama and H. Hashimoto, *Appl. Surf. Sci.*, 2008, **255**, 1162–1164.

1199 189 J. Seuma, J. Bunch, A. Cox, C. McLeod, J. Bell and C. Murray, *Proteomics*, 2008, **8**, 3775–84.

1200 190 J. S. Becker, A. Matusch, C. Palm, D. Salber, K. A. Morton and J. S. Becker, *Metallomics*, 2010, **2**, 104–
1201 111.

1202 191 C. Giesen, T. Mairinger, L. Khoury, L. Waentig, N. Jakubowski and U. Panne, *Anal. Chem.*, 2011, **83**,
1203 8177–8183.

1204 192 C. Giesen, H. A. O. Wang, D. Schapiro, N. Zivanovic, A. Jacobs, B. Hattendorf, P. J. Schöffler, D.
1205 Grolimund, J. M. Buhmann, S. Brandt, Z. Varga, P. J. Wild, D. Günther and B. Bodenmiller, *Nat.*
1206 *Methods*, 2014, **11**, 417–422.

1207 193 J. S. Becker, A. Matusch and B. Wu, *Anal. Chim. Acta*, 2014, **835**, 1–18.

1208 194 J. Olejnik, E. Krzymańska-Olejnik and K. J. Rothschild, *Methods Enzymol.*, 1998, **291**, 135–154.

1209 195 G. Thiery, R. L. Mernaght, H. Yan, J. M. Spraggins, J. Yang, F. F. Parl and R. M. Caprioli, *J. Am. Soc.*
1210 *Mass Spectrom.*, 2012, **23**, 1689–1696.

1211 196 J. Franck, K. Arafah, M. Elayed, D. Bonnel, D. Vergara, A. Jacquet, D. Vinatier, M. Wisztorski, R. Day, I.
1212 Fournier and M. Salzter, *Mol. Cell. Proteomics*, 2009, **8**, 2023–2033.

1213 197 and I. F. J. Stauber, R. Lemaire, M. Wisztorski, S. Ait-Menguellet, J. P. Lucot, D. Vinatier, A. Desmond,
1214 M. Deschamps, G. Proess, I. Rudlof, M. Salzter, in *HUPO 5th Annual World Congress*, p. 2006.

1215 198 G. Thiery, M. S. Shchepinov, E. M. Southern, A. Audebourg, V. Audard, B. Terris and I. G. Gut, *Rapid*
1216 *Commun. Mass Spectrom.*, 2007, **21**, 823–829.

1217 199 H. M. Santos and J. L. Capelo, *J. Proteomics*, 2012, **75**, 4921–4930.

1218 200 M. S. Shchepinov and V. a Korshun, *Chem. Soc. Rev.*, 2003, **32**, 170–180.

1219 201 G. Thiery, E. Anselmi, A. Audebourg, E. Darii, M. Abarbri, B. Terris, J. C. Tabet and I. G. Gut,
1220 *Proteomics*, 2008, **8**, 3725–3734.

1221 202 J. Yang, P. Chaurand, J. L. Norris, N. a Porter and R. M. Caprioli, *Anal. Chem.*, 2012, **84**, 3689–95.

1222 203 R. Hong, J. True and C. Bieniarz, *Anal. Chem.*, 2014, **86**, 1459–1467.

1223 204 M. van de Waterbeemd, K. L. Fort, D. Boll, M. Reinhardt-Szyba, A. Routh, A. Makarov and A. J. R.
1224 Heck, *Nat. Methods*, 2017, **14**, 283–286.

1225 205 H. Li, H. H. Nguyen, R. R. Ogorzalek Loo, I. D. G. Campuzano and J. A. Loo, *Nat. Chem.*, 2018, **10**, 139–
1226 148.

1227 206 L. S. Eberlin, X. Liu, C. R. Ferreira, S. Santagata, N. Y. R. Agar and R. G. Cooks, *Anal. Chem.*, 2011, **83**,
1228 8366–8371.

1229 207 H. Thiele, S. Heldmann, D. Trede, J. Strehlow, S. Wirtz, W. Dreher, J. Berger, J. Oetjen, J. H. Kobarg, B.
1230 Fischer and P. Maass, *Biochim. Biophys. Acta - Proteins Proteomics*, 2014, **1844**, 117–137.

1231 208 B. Heijs, W. M. Abdelmoula, S. Lou, I. H. Briaire-de Bruijn, J. Dijkstra, J. V. M. G. Bovée and L. A.
1232 McDonnell, *Anal. Chem.*, 2015, **87**, 11978–11983.

1233 209 W. M. Abdelmoula, R. J. Carreira, R. Shyti, B. Balluff, R. J. M. van Zeijl, E. A. Tolner, B. F. P. Lelieveldt,
1234 A. M. J. M. van den Maagdenberg, L. A. McDonnell and J. Dijkstra, *Anal. Chem.*, 2014, **86**, 3947–
1235 3954.

1236